

MATTINGLY, STANGER & MALUR, P.C.

JOHN R. MATTINGLY*
DANIEL J. STANGER
SHRINATH MALUR*

GENE W. STOCKMAN
Of Counsel

JEFFREY M. KETCHUM
SCOTT W. BRICKNER
Registered Patent Agents

* Bar Membership Other Than Virginia

ATTORNEYS AT LAW
104 EAST HUME AVENUE
ALEXANDRIA, VIRGINIA 22301

(703) 684-1120

PATENT, TRADEMARK
AND COPYRIGHT LAW
FACSIMILE (703) 684-1157

jc803 U.S. PTO
06/29/00

Date: June 29, 2000

Attorney Docket No. H-926

To: Assistant Commissioner for Patents
Washington, D.C. 20231

Sir: Transmitted herewith for filing is the patent application of:

Inventor: SEE ATTACHED LIST (S. HONMA et al)

For:
COMPUTER SYSTEM USING A STORAGE AREA NETWORK AND
METHOD OF HANDLING DATA IN THE COMPUTER SYSTEM

jc851 U.S. PTO
09/606050
06/29/00

Enclosed are:

☒

21 Sheets of Drawings

☒

This application is being filed without an executed Declaration.

Priority is claimed from _____ Application No. _____
filed _____ ☐ A certified copy is attached herewith.

Copies of the disclosure documents listed on the attached PTO 1449 form and
☐ discussed in the specification or ☐ attached Information Disclosure Statement.

A verified statement to establish small entity status under 37 CFR 1.9 and 1.27.

JAPANESE LANGUAGE APPLICATION

☒

Specification: Abstract ☒, Description 29 pages; and 12 claim(s).

☐

Preliminary Amendment.

☐

Executed Declaration.

The filing fee is calculated as shown below:

Small Entity

Large Entity

| For: | No. Filed | No. Extra |
|---|-----------|-----------|
| Basic Fee | | |
| Total Claims | 12 - 20 = | * 0 |
| Indep Claims | 3 - 3 = | * 0 |
| <input type="checkbox"/> Multiple Dependent Claim (s) | | |

* If difference is less than zero
then enter '0' in second column

| Rate | Fee |
|-------|--------|
| | \$ 345 |
| x 9 | \$ |
| x 39 | \$ |
| + 130 | \$ |
| Total | \$ |

OR

| Rate | Fee |
|-------|--------|
| | \$ 690 |
| x 18 | \$ 0 |
| x 78 | \$ 0 |
| + 260 | \$ 0 |
| Total | \$ 690 |

☒

A check in the amount of \$ 690.00 is enclosed for the filing fee.

☒

The Commissioner is hereby authorized to charge any additional fees that may be required to
Deposit Account No. 50-1417.

Respectfully Submitted,

By: John R. Mattingly
John R. Mattingly
Registration No. 30,293

09606050-062900

Fig. 1

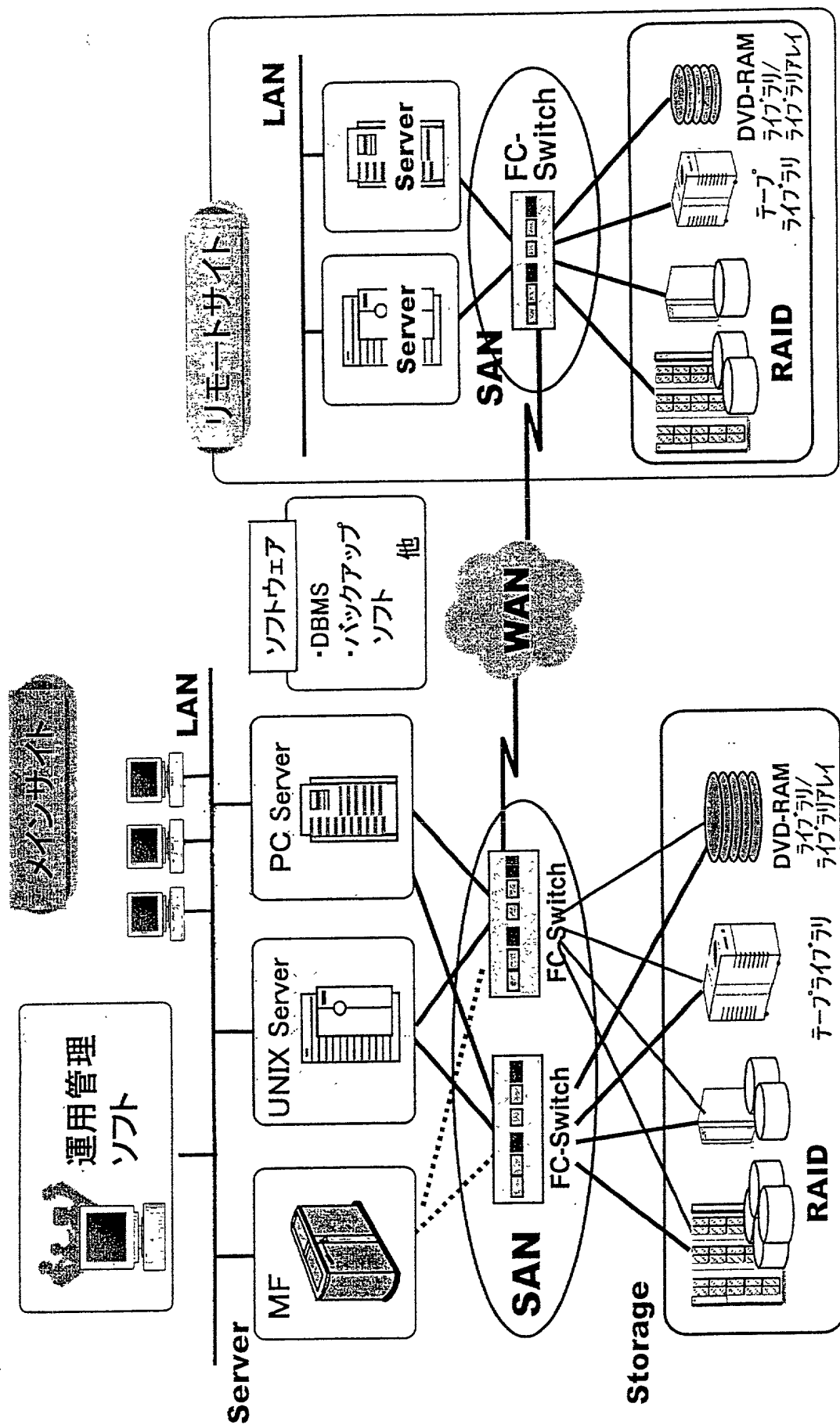


Fig. 2

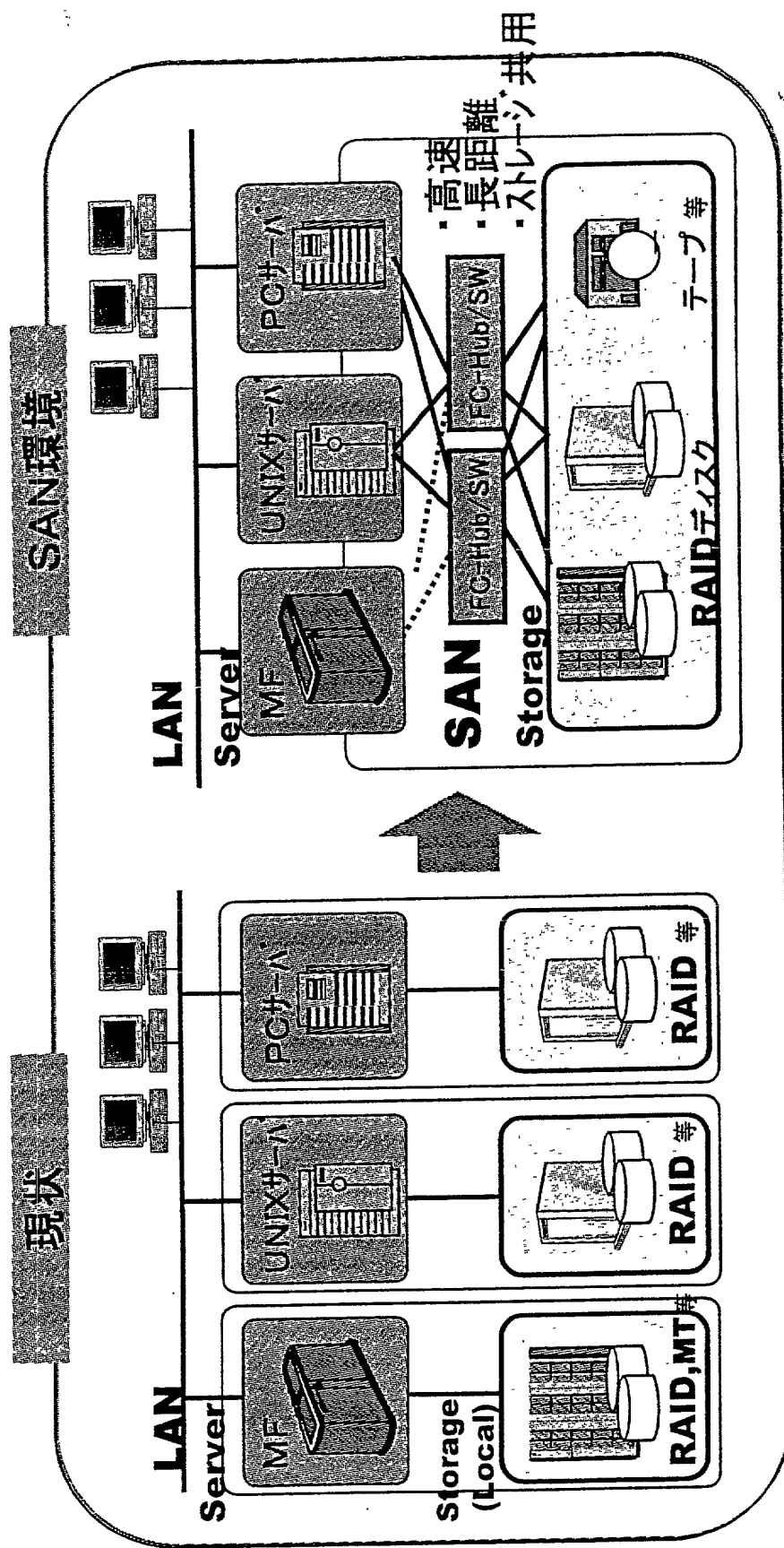


Fig. 3

統合ストレージシステムの主機能**Data Protection (障害対策、災害対策)**

- ・バックアップ(ディスクの障害に対する対策)
- ・ディザスタリカバリ(地震・火災等の災害に対する対策)

Data Sharing (データ共用)

- ・メインフレーム、UNIXサーバ、NTサーバからのデータ交換/共用
- ・多種情報の扱い(DB、文書、図面、マルチメディアコンテンツ等)

Storage Management (ストレージソース管理)

- ・複数部門で管理・運用しているストレージを一元管理
- ・統一的操作による環境設定、ストレージ運用管理

Fig. 4

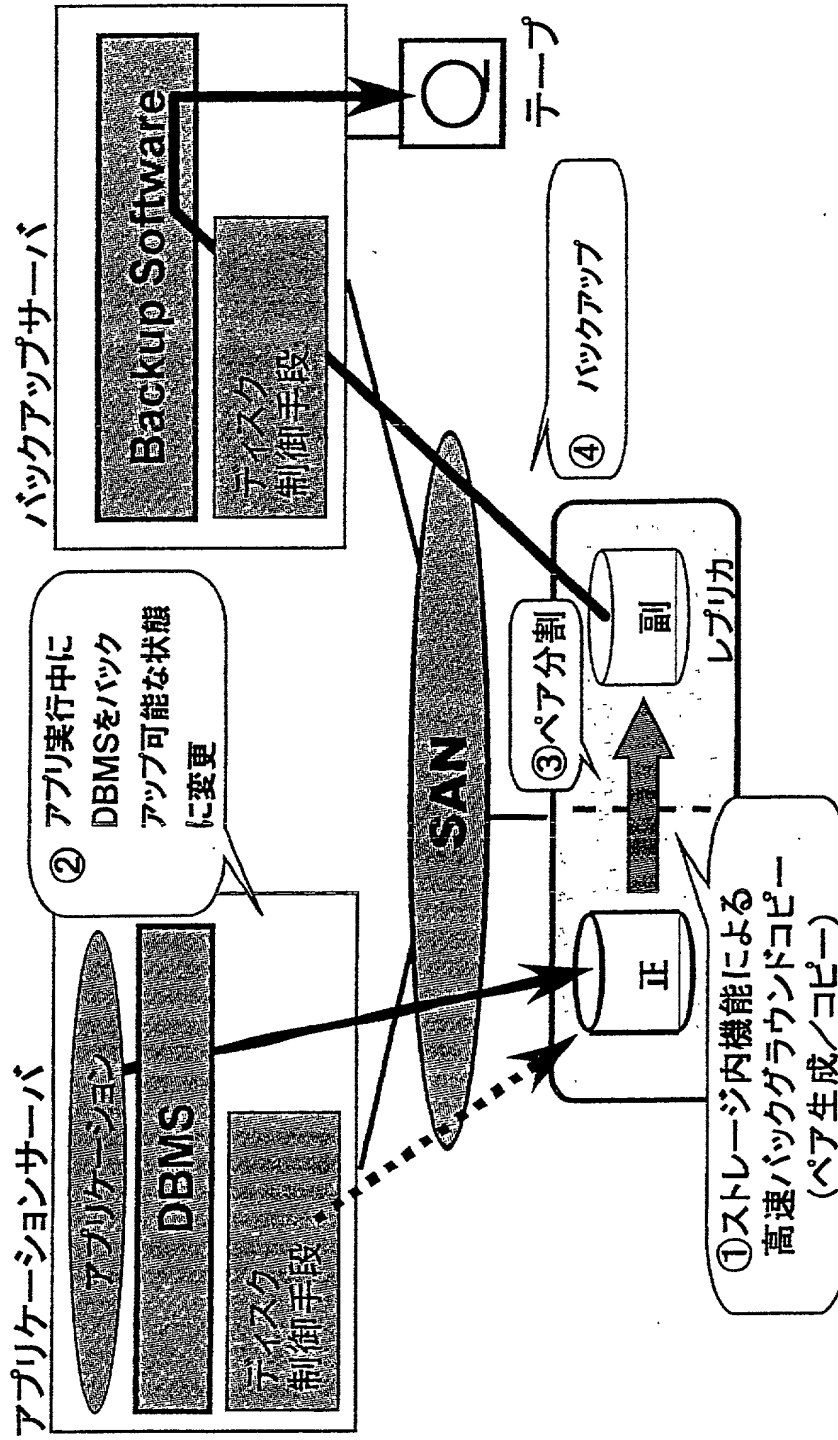
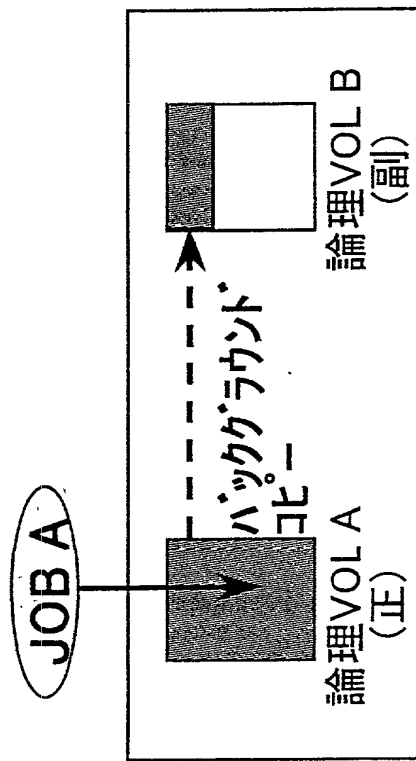
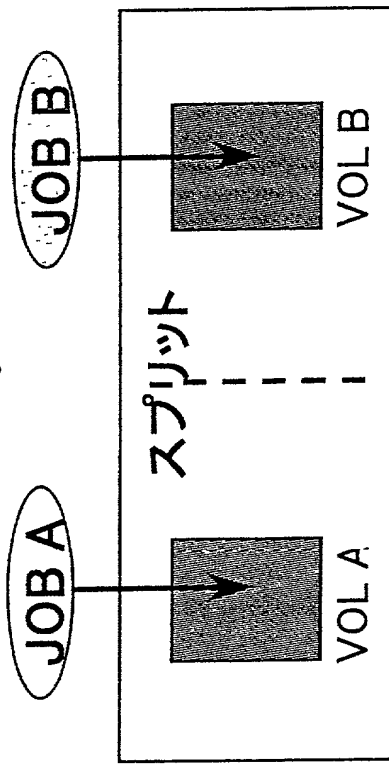


Fig. 5a

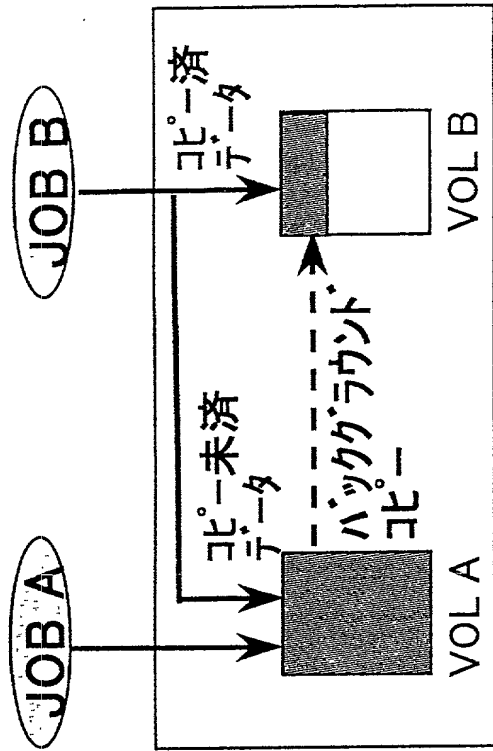


● 事前にバックグラウンドにてコピーを作成



● コピー完了後に副ボリュームをスプリットし別JOBから使用可能

Fig. 5b



- 見かけ上 副ボリュームを即スプリットし別JOBから使用可能
- 実際のデータはバックグラウンドで正から副へコピー

LIST OF INVENTORS' NAMES AND ADDRESSES

Shigeo HONMA, Odawara, JAPAN;

Hiroshi MORISHIMA, Yokohama, JAPAN;

Tokuhiro TSUKIYAMA, Nakagun, JAPAN;

Hiroyuki MATSUSHIMA, Yokohama, JAPAN;

Takashi OEDA, Sagamihara, JAPAN;

Yoji TOMONO, Hiratsuka, JAPAN.

006290-05090950

Japanese Language Patent Application

Title of the Invention

COMPUTER SYSTEM USING A STORAGE AREA NETWORK AND
METHOD OF HANDLING DATA IN THE COMPUTER SYSTEM

Inventors

Shigeo HONMA,
Hiroshi MORISHIMA,
Tokuhiro TSUKIYAMA,
Hiroyuki MATSUSHIMA,
Takashi OEDA,
Yoji TOMONO.

0360050 062900

ストレージエリアネットワークを用いた計算機システム及びそのデータ取り扱い方法

Computer system using a storage area network and method of handling data in the
computer system

0066050.062900

発明の背景

本発明は、データを記憶するストレージシステムに関し、特に、取り扱うデータのデータプロテクション、データ共用、ストレージリソース管理、並びにデータの取り扱い方法等に関する技術である。

現在、情報処理の扱う環境は、インターネット及びイントラネットの進展、データウェアハウス、電子商取引、情報サービス等のアプリケーションの拡大により、急激に変化してきており、それにともない取り扱われるデータ量は急激に増大している。

例えば、過去5年間でCPU性能が100倍向上しているのに対して、ディスク入出力性能の向上は10倍程度に留まっている。すなわち、急増するトラヒックに対する入出力性能の限界が危惧されるようになってきている。また、膨大なデータを処理するERP (Enterprise Resource Planning)、データウェアハウスなどのアプリケーションの普及、扱う情報(文書、図面、ビジュアルコンテンツ等)の多様化・マルチメディア化に伴い、年平均2倍の伸びで企業でのディスク容量が必要とされてきている。更に、企業等で使用されるストレージ容量の増大や、使われ方の多様化にともない、そのストレージ運用コストも増大している。更にまた、メインフレーム基幹データを各部署で共用し、活用されるようになってきている。

取り扱うデータ量の増大に伴う情報処理環境の状況を図2を用いて説明すると、サーバとストレージの関係は、例えば、大型コンピュータ用のサーバとしてMF (メインフレーム)、中型コンピュータ用のサーバとしてUNIXサーバ、小型コンピュータ用のサーバとしてPCサーバ、がそれぞれに専用のストレージ、例えばRAID (Redundant Arrays of Inexpensive Disks)、MT (Magnetic Tape) 等と接続されており、クライアントコンピュータはLANを通してそれぞれのサーバに指示を与えて、当該サーバに専用のストレージを用いてデータ処理を実行していた。

最近では、上述した各種サーバとストレージの間にSAN (Storage Ar

ea Network)を構築して、各種サーバはいずれのストレージにもアクセスでき得るSAN環境が提案されている。ここで、SANとは複数のサーバとストレージをファイバチャネルで接続したストレージ入出力専用ネットワークを意味しており、このネットワークの採用によって、各種ストレージを共用化でき、サーバとストレージ間を高速でデータ処理でき、更に長距離化できるようになっている。

発明の概要

上述したように、情報処理を扱う環境では、入出力性能の向上、ディスク容量、ストレージ運用、データ共用を進めるためにSANの導入が進められている。SANとは、図2に示すように、複数のサーバとストレージを高速のネットワーク（例えばファイバチャネル）で接続した新しいネットワークである。この環境においては、サーバ毎に接続されて各サーバに従属していたストレージを独立化させ、ストレージ専用のSANネットワークがまず構築される。また、SANネットワーク上のストレージ情報をアクセス権のあるすべてのユーザから共有させることができる。

また、複数のストレージを接続するためストレージ入出力性能を大幅に向上することができ、即ち、メリットとして、ストレージ入出力性能の大幅向上（パフォーマンス向上）、サーバ環境とは独立した柔軟なストレージ環境の設定・拡張（スケーラビリティの向上）、ストレージの一元運用（ストレージ管理機能の向上）、接続距離を飛躍的に延ばす事による災害対策（データ保護機能の向上）等を図ることができるようになった。

しかしながら、提案されているSANネットワークでは、これらを実現するための具体的な構成、又は実施方法等は必ずしも明確に開示されていなかった。

本発明の目的は、SANの採用による種々の利点並びに有用性を確保できるように、ストレージシステムの具体的機能並びにこれらの機能に対応する具体的構成を工夫してストレージシステム全体の連携を強化した統合ストレージシステムを提供するとともに、インターネットに接続して大容量のデータ保存とその利用を図るインターネッ

トデータセンタ（「iDC」と以下称する）へ統合ストレージシステムを適用し、iDCでのより有用なデータ取り扱い方法を提供することにある。

前記課題を解決するために、本発明は主として次のような構成を採用する。

複数のクライアントコンピュータ、複数の各種サーバ、データを保存する複数の各種ストレージ、前記コンピュータと前記サーバを接続するローカルエリアネットワーク（LAN）、前記サーバと前記ストレージの間に介在するストレージエリアネットワーク（SAN）、を備えた計算機システムであって、前記SANは、ファイバチャネルスイッチ（FCスイッチ）によって前記サーバと前記ストレージを任意に相互接続し得る回線網を形成し、前記各種ストレージの論理ボリューム、データ配置、障害監視を含むストレージ管理と、前記FCスイッチの設定管理と、前記ストレージのデータのバックアップ運用と、を行う管理運用ソフトウェアを備えた端末を有するようにする。

また、サーバ、前記サーバのデータを記録するストレージ、前記サーバと前記ストレージを接続するネットワークから構成されるシステムの管理方法であって、処理対象であるデータを識別する情報を入手し、前記情報が示すデータの処理内容を入手し、前記情報が示すデータを保持する前記ストレージに対して、前記処理内容を指示し、および、前記情報が示すデータが処理された結果を前記ストレージから受けるようにする。

図面の簡単な説明

図1は、本発明の実施形態に係る統合ストレージシステムの基本的な全体構成を示した概略図である。

図2は、従来技術のストレージシステムの全体構成を示した概略図である。

図3は、本発明の実施形態に係る統合ストレージシステムの主たる機能を説明する図である。

図4は、本実施形態の無停止バックアップについての基本的なシステム構成を示す

図である。

図 5 a、図 5 b は、本実施形態の無停止バックアップについての機能乃至作用を説明する図である。

図 6 は、本実施形態の無停止バックアップについてのミラーリングソフトを用いたシステム構成を示す図である。

図 7 は、バックアップシステムにおける事前準備とシステム構築の例を示す図である。

図 8 は、本発明の実施形態に係るテープを共用するバックアップの種々のシステム構成例を示す図である。

図 9 は、複数サーバが 1 台のテープライブラリを共有するテープ共用バックアップの構成を示す図である。

図 10 は、本発明の実施形態に係るディザスタリカバリでの非同期リモートコピーのシステム構成を示す図である。

図 11 は、本発明の実施形態に係るデータシェアリングでのサーバ間 DB 高速レプリケーションのシステム構成を示す図である。

図 12 は、本発明の実施形態に係るシステム統合運用管理の内の障害監視とバックアップ運用を示す図である。

図 13 は、本発明の実施形態に係るシステム統合運用管理の内のストレージ性能の集中管理を示す図である。

図 14 は、本発明の実施形態に係るシステム統合運用管理の内のストレージ管理、特に LUN マネージャと LUN セキュリティを示す図である。

図 15 は、本発明の実施形態に係るシステム統合運用管理の内のストレージ管理、特にサブシステム内の階層制御を示す図である。

図 16 は、本発明の実施形態に係るシステム統合運用管理の内のスイッチ管理、特にゾーニング設定を示す図である。

図 17 は、本発明の実施形態に係る統合ストレージシステムを用いたインターネッ

トデータセンタのシステム構成の概要を示す図である。

図18は、本実施形態のインターネットデータセンタにおけるストレージ統合を示す図である。

図19は、本実施形態のインターネットデータセンタにおける無停止バックアップのシステム構成を示す図である。

図20は、本実施形態のインターネットデータセンタにおけるセキュリティ確保のシステム構成を示す図である。

図21は、複数の企業の計算機システムが相互結合された大規模計算機システムの構成例を示す図である。

実施例の詳細説明

本発明の実施形態に係る、ストレージエリアネットワーク（SAN）を用いた計算機システム及びそのデータ取り扱い方法について、図面を用いて以下説明する。図1は、本発明の実施形態に係る前記計算機システムの基本的な全体構成を示した概略図である。

図1において、SANを用いた計算機システムは、メインサイトとリモートサイトとから成り立っていてこれらのサイトはWAN（Wide Area Network）で接続されている。メインサイトにおいては、複数のクライアントコンピュータと各種サーバ、例えば、大型コンピュータ用サーバのMF（メインフレーム）、中型コンピュータ用サーバのUNIXサーバ、小型用コンピュータ用サーバのPCサーバ（パソコン用サーバ）がLANで接続されている。また、LANには統合ストレージシステムの運用管理ソフトを搭載した専用端末が接続されていてこの総合ストレージシステム全体が運用・管理・監視されるようになっている。なお、この運用管理ソフトは専用端末ではなくてどれか1つのクライアント端末に搭載されていて、当該クライアント端末が統合ストレージシステムの運用管理を行っても良い。

更に、RAID、テープライブラリ、DVD-RAMライブラリ／ライブラリアレ

イ等のストレージは、メインフレーム（MF）サーバ、UNIXサーバ、PCサーバ等のサーバと、ファイバーチャネルスイッチ（FC-Switch）、不図示のファイバチャネルハブ（FC-Hub）等のネットワークスイッチで構成されるSAN（Storage Area Network）で接続される。また、メインサイトはWAN等の広域通信網により、メインサイトと同様な構成を持つリモートサイトに接続される。

ここで、サーバとストレージはSANにおけるチャネルスイッチにより接続されるので、チャネルスイッチに接続されるサーバやストレージを自由に追加、削除、変更することが可能となる。このため、まず、ストレージ容量や、記憶するデータの種別や目的（アクセススピード、コスト等）に合致したストレージを自由に追加、削除等が可能となる。また、サーバ側でも、チャネルスイッチを介して自由にこれらのストレージにアクセス可能となる。

また、WANを介して、リモートサイトと接続されているので、サイト間でデータの共用が可能となり、世界的規模で大量データの共用が可能となる。また、メインサイト及びリモートサイト間でデータのコピーを保持すれば、災害等で片方のサイトが障害になった場合でも他方のサイトのデータを用いて、業務を継続できることが可能となる。この場合、リモート側のサイトでのバックアップは必ずしも同一形態の記憶装置、例えばメイン側のRAIDからリモート側のRAIDへのコピーに限定されず、メイン側のRAIDからリモート側のDVD-RAMやテープライブラリ等へコピーすることで、コストや管理を簡単にすることも考えられる。この場合、SANの管理用端末の運用管理ソフトがこれらデータのコピー元、コピー先等を管理する。

また、図2に示す従来技術では、クライアントは、その用途別に例えばメインフレーム、UNIXサーバ、PCサーバとLAN等の通信回線を介して、また、各サーバ間もLANを介して接続されていた。そして、それぞれのストレージは各々のサーバに接続されていた。このため、ストレージに格納されているデータは各サーバ経由でしかアクセスできなかった。

一方、本発明の実施形態では、各サーバに接続されるストレージに格納されるデータをS A Nにより統合的に管理する。まず、複数の各サーバからS A Nを構成するファイバチャネルスイッチ（F C - S w i t c h）を介して、各種ストレージ（R A I Dディスク、テープライブラリ、D V D - R A Mライブラリ／ライブラリアレイ等）と接続する。これにより、各ストレージが有するデータはL A Nを経由せずに、各サーバから直接アクセス可能となり、例えば、大量データ等のアクセスが容易となる。また、データ用のストレージが統合ストレージシステムに整理統合されるので、データや機器の管理が容易となる。

また、従来技術では、災害に対するデータのバックアップやリモートコピー等は各サーバ毎に対応するストレージを設置しL A N経由で実行することが必要であったが、本発明の実施形態ではS A N並びに各種ストレージからなる統合ストレージシステムを設けるので、この統合ストレージシステムでデータのバックアップが可能となり、より効率的なデータのリモートバックアップが可能となる。

以上、概略的に説明したS A N適用による計算機システムは、取り扱うデータについて、いつでも、誰でも、どこからでも、どんな情報でも安全に利用できることを主眼とした情報システムであることが求められるのである。

そして、本発明の実施形態に係る統合ストレージシステムは、図3に開示するように、ディスクの障害に対する対策としてのバックアップや地震・火災等の災害に対する対策としてのディザスタリカバリを内容とするデータプロテクションを基本的機能の一つとして持ち、また、メインフレーム、U N I Xサーバ、P Cサーバからのデータ交換・共用やD B（データベース）、文書、図面、マルチメディアコンテンツ等の多種情報の扱うデータシェアリング（D a t a S h a r i n g：データ共用）を基本的機能の一つとして持ち、更に、サーバ毎に管理運用していたストレージの一元管理や統一的操作による環境設定並びにストレージ運用管理を内容とするストレージマネージメント（ストレージリソース管理）を基本的機能の一つとして持つ。

以下、順に本発明であるそれぞれの基本的機能の詳細な具体的内容について説明す

る。これら機能は、これら機能を記述したプログラム（ソフトウェア）や必要なデータを、ストレージ、スイッチ、サーバ（コンピュータ）、および管理装置（コンピュータ等で実現される）等の装置が有するメモリにインストールし、各装置が有するCPU（中央演算装置）で実行することによって実現される。なお、SANを適用した大容量のストレージシステム群と各種サーバとからなる計算機システムをインターネットに接続してデータ保存サービス機能を具備させたデータセンタ、即ちインターネットデータセンタ（以下、「IDC」と略称する）を構築し、IDCにおける膨大なデータの取り扱い方法に関する工夫も本発明の特徴の一つである。

まず、データプロテクションについてであるが、その機能は、オンライン稼働中のDBバックアップ、ストレージ資源共用による管理費節減、ディザスタリカバリ等によるシステム可用性の向上、及びデータの安全性確保を狙ったものであり、今後増加が予想される24時間365日の営業に対してでも、業務を停止すること無くデータのバックアップ（無停止バックアップ）が可能となると共に、バックアップに際してテープライブラリ等のバックアップ装置を共用可能とすること（テープ共用バックアップ）でコスト削減が可能となり、更に、遠距離リモートコピーにおけるデータ安全性を確保して災害時にシステムを迅速に復旧すること（リモートコピー）が可能となる。具体的には、データプロテクションの内容は、前述したように、無停止バックアップとテープ共用バックアップと非同期リモートコピーの3つの技術である。

まず、無停止バックアップの機能乃至作用は、データのレプリカからのバックアップによりバックアップ中のアプリケーションの動作を可能とするとともに、バックアップ専用サーバにより業務用サーバに影響を与えないようにすることである。

無停止バックアップの構成と機能について、図4と図5a、図5bにその詳細を示す。この機能の概要は、ストレージ内に存する内部機能とアプリケーションサーバ内のDBMS（データベースマネジメントシステム）との連携により、オンライン業務に影響を与えずに、LAN非経由且つSAN経由でDBをバックアップすることである。

図4では無停止バックアップの一連の流れを示している。まず、前記ストレージ内の内部機能によって、ストレージ装置内でバックアップ対象ボリューム（正ボリューム）と同一あるいはそれよりも大きい容量を持つ副ボリュームに対してコピーを実行して正ボリュームのコピーを作成する。続いてアプリケーション実行中に、アプリケーションサーバ内のDBMS（データベースマネジメントシステム）をバックアップ可能な状態に変更してオンライン業務に影響を与えないようにしておき、副ボリュームの内容をバックアップサーバによりテープにバックアップする。

図5a、図5bには図4の無停止バックアップにおける、ストレージ装置の内部機能であるボリュームコピー機能の処理概要を示している。もともと、不図示の従来のバックアップ技術ではDB（データベース）へのサーバからの業務を停止したあとに、DBを他のストレージにバックアップして、当該バックアップ完了後に前記DBへのオンライン業務を再開するというものであった。この従来技術によると、バックアップ中はDBへのオンライン業務は停止せざるを得なかった。

これに対して、図5aに図示した本実施形態の一構成例では、ストレージ内にバックアップ用のレプリカ、すなわち論理ボリュームB（論理VOLB）を確保して事前にコピーを作成しておくものである。論理ボリュームA（論理VOLA）のデータバックアップの場合、事前に論理VOLAの内容を論理VOLBの方にもコピーしておく。具体的には、論理VOLAがバックアップ対象である場合には、あらかじめ論理VOLAと論理VOLBの2つのボリュームを用意し、二重化を指示する。

ストレージ装置では、論理VOLAの内容を論理VOLBに順次コピーしていき、それと並行してオンライン業務（図中JOBA）からの書き込みがあると、ストレージ装置内で自動的に論理VOLA、論理VOLBの双方に二重書きする。論理VOLAから論理VOLBへの順次コピーが完了した後も、JOBAからの書き込みがあれば二重書き込みは実施し、論理VOLAと論理VOLBの内容を常に同一に保つ。

バックアップを実施する際には、バックアップサーバはディスク制御手段を用いてストレージ装置にペア分割（スプリット）を指示する。ストレージ装置ではスプリッ

ト指示のあとは、J O B Aからの書き込みがあっても論理VOL Aのみに書き込みし、論理VOL Bには書き込みを実施しない。これにより論理VOL Bにはスプリット指示のあった時点の論理VOL Aの内容がそのまま残される。スプリット指示の後、バックアップサーバのバックアップソフトウェアは副ボリュームである論理VOL Bからデータを読み出し、テープ装置などのバックアップデバイスにデータのバックアップを行う。

ただし、図5 aによるボリューム二重化の方式では、バックアップを実施するよりも前の時点で二重化ボリュームを用意しておく必要がある。そのため、バックアップを実施するには、事前にボリューム二重化にかかる時間を見越して、それより前に二重化を行う必要がある。図5 bに図示されたストレージ装置の機能では、この問題を解決する。

図5 bの事例でも論理VOL Aのコピーを置くための論理VOL Bを用意する点は図5 aと同様である。バックアップサーバは、図5 aの例と同様、バックアップを開始する前にディスク制御手段を用いてストレージ装置にペア分割（スプリット）を指示する。ただし、この時点で論理VOL Bには論理VOL Aのデータがコピーされている必要はない。スプリット指示の後、バックアップサーバのバックアップソフトウェアは、副ボリュームである論理VOL Bからデータの読み出しを開始する。ストレージ装置では論理VOL Aから論理VOL Bに対して、順次データのコピーが行われるが、バックアップサーバが論理VOL Bからデータを読み出そうとしたときにボリュームBにデータが存在しない場合には、ディスク装置は論理VOL Aからデータを読み出してバックアップサーバに渡す、あるいは論理VOL Aから一旦データを論理VOL Bにコピーし、それをバックアップサーバに渡す。この処理により、スプリット時点で論理VOL Bにデータが存在していなくとも、バックアップサーバからはあたかも論理VOL Bに論理VOL Aのコピーが存在するように見える。

ただし、バックアップ処理の間、アプリケーションサーバから論理VOL Aのある領域にデータが書き込まれることもあり得る。ストレージ装置では論理VOL Aから

論理VOL Bに対して、データを順次コピーしているので、もしもこのデータがこのコピー処理によって論理VOL Bに書き込まれてしまうと、論理VOL Bにはスプリット以降のデータが書き込まれてしまうことになる。これを防ぐため、ストレージ装置は論理VOL Aの書き込み要求のあった領域に現在書かれているデータを読み出し、論理VOL Bへと書き出す。しかる後に、アプリケーションサーバから書き込み要求のあったデータを論理VOL Aに書き込む。この処理により、論理VOL Bにはスプリット指示のあった時点の論理VOL Aのデータ内容だけがコピーされることになる。この方法では、バックアップ開始時点で副ボリューム（論理VOL B）に正ボリューム（論理VOL A）の内容がコピーされている必要がない、すなわち事前にコピーボリュームを用意するという運用が不要になるので、運用性が向上する。

図7は、図4と図5 a、図5 bに示す無停止バックアップの際のシステム構築のインストール事例を示す。アプリケーションサーバにはDBMSとディスク制御手段が備えられ、バックアップサーバにはバックアップソフトとディスク制御手段が備えられている。事前準備として、ディスク制御手段をインストールしてその構成を設定し動作確認を行う。その後、無停止バックアップシステムの構築に際しては、まず、DBMSのスク립トを作成し（ログイン、バックアップモード設定、バックアップモード解除、ログアウト）、アプリケーションサーバにおけるディスク制御手段のスク립トを作成し（ペアスプリット、ペアイベント待機、再同期）、バックアップソフトとの連携動作を確認し、Logical Unitの配置とディスク制御手段のパラメータ設定をする。

また、図6に示す無停止バックアップの他の構成例によると、ミラーリングソフトウェアにより形成されたプライマリ（正）とセカンダリ（副）のボリュームを、アプリケーションサーバの連携ツールの指示によりミラースプリットし、一方（副ボリューム）でバックアップを行うとともに、他方（正ボリューム）で業務を継続可能とするものである。そして、バックアップ終了後には再同期を行う。具体的には、アプリケーションサーバ内のミラーリングソフトにより正と副のボリュームに二重書きを行

い、サーバ内の連携ツール（ソフトウェア）でDBアクセスを停止し、ミラーズブリットを指示した後にDBアクセスを再開する。次いで、バックアップサーバ内の連携ツール（ソフトウェア）で副ボリュームのデータをバックアップサーバに接続されたテープ等のバックアップ装置にバックアップを開始する。その後、バックアップサーバ内の連携ツール（ソフトウェア）からのバックアップ完了を受けたアプリケーションサーバ内の連携ツールは、ミラーの再同期を指示して再び二重書きする。

次に、テープ共用バックアップの構成と機能について、図8と図9にその詳細を示す。この機能の概要は、多数のサーバに点在するデータの管理にかかる管理コストを低減するとともに、LANへの負荷を低減した高速のバックアップを図るものである。そして、テープライブラリを複数サーバ側から共用可能とすることで高価なライブブラリを有効に使用可能とし（個々のディスク毎にバックアップテープを設けるもの比べて）、単一のテープライブラリを複数のサーバで共同、共有することでLAN経由ではなくてSAN経由で直接にテープ装置に出力できてバックアップを高速化できる。

図8の左図は従来のテープバックアップを示したものであり、サーバ毎のディスクからLANを経由してバックアップサーバを通してテープにデータをバックアップしていて、バックアップ毎にLANを経由するためLANへの負荷が掛かるものであり、更に、バックアップサーバにもバックアップ毎に負荷がかかるものであった。

本発明の実施形態は、図8の中央図のLANフリーバックアップによると、SAN経由でディスクからテープへコピーし高速化することができ、LANを経由することなくサーバを用いてバックアップするものである。バックアップに際してサーバは単一のものを使用することもできるので、サーバへの負荷も低減する。本発明の他の実施形態は、図8の右図のサーバレスバックアップによると、ディスクから直接にテープにコピーすることができるのでバックアップの高速化とサーバの負荷低減ができる。図8の右図の実施形態では、ディスクにテープへの書き込み機能を具備させる必要があり、又はテープにディスク内のデータ読み出し機能を具備させる必要があり、又はFCスイッチにディスクからテープへの書き込み機能を具備させる必要があり、又は

テープがFC-SCSIマルチプレクサ（図9の説明で後述する）に接続される場合にFC-SCSIマルチプレクサがディスクからテープへの書き込み機能を具備させる必要がある。

テープ共用バックアップの他の構成例を図9に示す。図9の構成は図8の中央図に示すLANフリーバックアップに対応するものである。この構成例では2ノード以上で同時にテープライブラリを共有して各々のサーバがバックアップするものである。図9によると、サーバCはサーバA及びBとその機能を異にしており、バックアップを実際に執行するに必要なバックアップエージェントの外に、バックアップを統合管理するバックアップマネージャを保持していて、バックアップドライブを割り当てる機能等を備えている。ここで、バックアップドライブは、例えば、ドライブを3個有していてサーバAにはドライブ1を割り当て、サーバAからバックアップしたいという要求があれば保存用テープをドライブAにローディングするように制御される。また、サーバに対するドライブの割り当ては、バックアップマネージャがドライブの使用状況を管理し、未使用のドライブを選定してその内の適宜のドライブを割り当てても良い。図9に示す構造において、FC-SCSIマルチプレクサとバックアップドライブとが図8に示すテープライブラリに相当する。

図9に示すテープ共用バックアップの具体的動作を説明する。まず、サーバAのエージェントがバックアップマネージャに対してテープマウントの要求を行う。次に、マネージャが当該要求を受けてテープライブラリのドライブのどれか1つにテープをマウントする。続いて、マネージャがサーバAのエージェントに対してマウント完了とマウントしたドライブ名を教える。次に、サーバAのエージェントが実際にバックアップを実行する。具体的には、サーバAがストレージからデータを読み出して、FCスイッチ、FC-SCSIマルチプレクサを介してマウントしたテープに書き込む。続いて、バックアップ完了後にサーバAのエージェントはマネージャに対してテープのアンマウント要求を行い、マネージャはテープアンマウントを指示して、全てを終了する。

次に、データプロテクションの1つとしてのディザスタリカバリにおける非同期リモートコピーについてその構成と機能について説明する。これは、遠距離リモートコピーによってデータ保全性を確保して地震などの災害時にシステムを迅速に復旧させるものであり、メインサイトの性能への影響を無くしてリモートサイトへデータベースを複製し、災害時にリモートサイトで業務を続行しようとするものである。

図10は非同期リモートコピーのシステム構成を示しており、メインサイトとリモートサイトとは災害時に同時に被災しないだけの遠距離を通信回線を介して接続され、メインサイトで情報が更新されそれが完了すると（リモートサイトへの情報の反映するのを待たずに、即ち非同期で）、更新完了をサーバに報告する。次に、適宜のタイミングで更新データをメインサイトからリモートサイトへコピーするのであるが、メインサイトへの更新データの順序通りにデータ転送されない場合には、リモートサイトのシステムで更新データを時系列にソートして更新の順序性を保証してコピーしている（例えば、入金と出金の更新データが順序を逆にしてデータ保存されると残高処理で不都合な取り扱いをせざるを得ないという事態が発生し得る）。

次に、データシェアリングにおけるサーバ間高速レプリケーションの構成と機能について説明する。図11に示すように、メインフレームのDB（基幹データベースで高信頼性を確保できるもの）とUNIX/NTサーバのDB（例えば、データの信頼性よりもデータの統計処理を実行する上でデータの取り扱い易さを重視してメインフレームDBから統計処理の必要原データをロードするためのもの）間でデータをロードする際に、メインフレームベースのファイルである中間ファイルを立てて、一旦この中間ファイルに基幹DBからデータを落とし込む（UNIXサーバのデータローダーが基幹DBから直接にデータを読み込むような仕様とはなっていないので）。中間ファイルのデータはUNIXサーバのデータローダーで読み込みが可能なようレベルに変換されているのでパイプを通してUNIXサーバのDBにレプリケーションを作成して所要の処理のためのDBとする。この際に、基幹DBからUNIXサーバのDBへのデータレプリケーションにおいてLANを経由していないためにサーバ間DB

006690.050960
の高速レプリケーションが可能となる。ここで、中間ファイルは、磁気ディスクの外に半導体メモリ、即ちキャッシュメモリ上に一時的に形成された仮想ボリュームであっても良い。キャッシュメモリを用いる方が高速にデータを転送できる。

更に、UNIXサーバ又はPCサーバのデータウェアハウス構築のし易さのために、基幹DB等の各種ソースDBからのデータ抽出、データの変換・集約、ロードの一連の処理をGUIベースに簡単・迅速に実行するソフトウェアをUNIXサーバ内に又はそれに付属させて設けることで、データウェアハウス構築時のデータ転送処理時間を短縮する事が可能となる。

次に、ストレージを含むシステム統合運用管理の構成と機能について以下説明する。大規模でかつ24時間連続運用を要求される計算機システムになると、システムの管理、とくにストレージの管理が重要になる。

ストレージの管理としては、装置障害、とくに装置内のどの部品が壊れたかを障害監視する機能が代表として挙げられる。また、システムクラッシュに備えて各サイトのデータを定期的にバックアップするなどの保守作業、ボリュームの追加を行った場合にシステム設定の変更作業、さらに、特定のボリュームに負荷が集中して性能が低下した場合にはあるボリュームのデータを別ボリュームに移動させるなどの措置が必要になり、その際の負荷状況の監視も重要な管理作業である。従来のシステムでは、各ストレージ装置ごとに1つの保守端末が存在し、各ストレージの管理をそれぞれ別の端末から行う必要があった。

本発明の実施形態に係るストレージ統合運用管理手段では、すべての装置を1箇所の端末から管理することができる。

図12は、大規模なオフィスシステムにおける、バックアップ運用・障害監視の事例を示している。一般のオフィス環境では各部門で共通に使用するデータ、全部門で共通に使用するデータとが存在する。本事例では各フロアA、B、Cごとに複数のクライアント計算機、複数のサーバ計算機が存在し、全社システムとして全部門で共通に使用されるメールサーバ、WWW (World Wide Web) サーバを用意

し、各部門に提供する。

部門ごとで使われるような小規模なデータの場合には、各部門内でバックアップを取れば済むケースも多いため、テープなどのバックアップデバイスは各部門に存在する。また、大規模なデータを格納するために複数の大規模ストレージとテープライブラリ等のバックアップ装置を計算機センタなどに用意し、センタの各デバイスと各フロア、全社システムとをストレージエリアネットワークにより相互結合する。

集中監視コンソールでは、各フロア、全社システム、計算機センタのすべての装置を監視し、装置障害の報告はすべて集中監視コンソールに集められる。保守員がコンソールを見てどの装置に障害が発生したか、容易に知ることができ、障害によりデータが壊れた場合には、バックアップ装置からデータを回復（リストア）することができる。このリストア処理も集中監視コンソールから可能である。また、保守員が端末から離れている場合もあるため、そういった場合には集中監視コンソールから保守員の持つ携帯電話などにメールを送信して保守員に知らせるなどの機能も存在する。

集中監視コンソールはバックアップの運用方法の指示・管理も行う。バックアップの頻度、バックアップ先の要求はそれぞれのデータの種類によって異なる。たとえば、ほとんどバックアップする必要のないデータ（例えばほとんど更新されないようなデータ）や特定の部門・個人しかアクセスしないようなデータを頻繁にバックアップする必要はない。あるいはすべてのデータを同じ時間帯にバックアップを取ろうとしても、バックアップデバイスの数には限りがある。集中監視コンソールでは、ユーザの必要に応じてそれぞれのデータあるいはボリュームによって、バックアップの頻度、時間帯、バックアップ先などを変え、それぞれのバックアップ処理を自動的に実施する。

図14は、ボリュームの設定を行う処理を模式的に示している。大規模なストレージ装置の場合、複数のディスクをまとめて1つないし複数の論理デバイス（LDEV）に見せかける。また、ホストまたはファイバチャネルスイッチに接続するための複数のポートを持ち、各LDEVがどのポートからアクセスされるようになるか設定変更

が可能である。ホストからLDEVを参照する場合、ストレージ装置のポート識別子とLUN (Logical Unit Number) により一意に認識できる。以下、これをホストアドレスと呼ぶ。ストレージ装置では、各LDEVにこのホストアドレスを割り付けてホストに見せる。

集中監視コンソールからは、LDEVのホストアドレス割付、そして、各LDEVにアクセスできるホストの種類を設定する。ストレージエリアネットワークではすべてのホストがすべてのストレージに接続されるため、本来アクセスしてはいけないホストが不当にストレージにアクセスする危険性があり、それを防ぐためにストレージ装置には、アクセス可能なホストの種類を登録することができる。

図13はストレージの性能監視の事例を示す。集中監視コンソールからは、各ボリュームの負荷状況を観測することができる。具体的には1秒間に何回のI/O処理を受け付けたか、リード・ライトの比率、キャッシュヒット率などである。一般にすべてのボリュームに均等に負荷がかかることは少なく、極端に負荷が高いボリュームやほとんど負荷がかからないボリュームとが存在することがある。集中監視コンソールでは複数のボリュームの負荷の偏りを一度にモニタできるので、これをみて、負荷の高い論理ボリュームのデータを一部低負荷のボリュームに再配置させるなど、システム全体の性能が落ちないように計画を立てることが容易になる。

また、図15は、ボリュームの再配置の機能をストレージ装置が持っている場合の事例を示したものである。ストレージ装置によっては、容量は少ないが比較的高速なボリューム、そして容量は大きいものの性能があまり高くないボリュームとが存在する。この場合、アクセス頻度の低いデータは容量の大きいボリュームに移動し、アクセス頻度の高いデータは高速なボリュームに移動することが望ましい。本事例のディスク装置では論理デバイス (LDEV) ごとに別の領域に移動することができる。

また、論理デバイスの移動中、移動後も、ホストからはボリューム配置が変わったことは見えず、移動前とまったく同じように扱うことができる。ディスク装置は論理デバイスの利用率などを統計情報として取得し、その情報を統合監視コンソールに送

る。統合監視コンソールではその情報を元に、ある論理デバイスを移動したときの、各論理デバイスの利用率がどう変化するかを予測し、保守員に提示する。保守員はその情報を元に、前図の例よりもより容易にボリューム再配置の計画を立てることができる。また、統合監視コンソールからは、実際に移動を実施するか否かを指示、あるいは、各ボリュームがある状態になったら自動的にボリュームを移動する、など細かい設定を事前に行うことなどが可能である。

また、システム統合運用管理の1つとしてFCスイッチの管理があり、FCスイッチの各種設定やゾーニング等の状態管理が可能となる。具体的には、fabric ポロジの表示、スイッチのゾーニングの設定、スイッチ内の各種パラメータの設定／表示等の管理があり、これらは総合監視コンソールで監視できるようになっている。

サーバとストレージに介在する fabric switch (FC) を3つのゾーニングに分割した構成例を図16に示す。

次に、以上説明した本発明の実施形態に係る計算機システムの全体構成の上で、図1に示した運用管理ソフトを搭載した端末である管理装置が計算機システム全体構成を管理制御する場合の具体的な例を示す。

バックアップ(図4)を行うには、ストレージ内のどのボリュームのデータをバックアップするかを決定する必要がある。通常、サーバはアプリケーションがストレージ内に記録したデータをファイル単位で管理しており、一方、ストレージはボリューム単位でデータを管理している。

従って、バックアップを開始するにあたり、SAN管理装置(図1に示す運用管理ソフト搭載の端末)は、サーバからあるファイルに関するバックアップを依頼されると、サーバからファイルを識別する情報や、バックアップ装置に関する情報(SAN上のアドレス、等)や、バックアップ時刻、等を手りする。SAN管理装置は、更に、当該ファイルを格納したボリュームを識別する情報をストレージから入手する。次に、SAN管理装置は、入手したこれら2つの情報を用いて、該ファイルを格納しているストレージに、バックアップの対象となるボリュームに関して、レプリカ(副ボリュ

ーム)を作成するよう指示する。具体的には、該ファイルを格納しているボリュームを有するストレージに、当該ボリューム(正ボリューム)のレプリカを作成するための別ボリューム(副ボリューム)の割り当てとレプリカの作成を指示する。副ボリュームの割り当ては、少なくとも正ボリュームと同容量のボリュームを割り当てるなどの配慮が必要であり、SAN管理装置は各ストレージがどういった容量、構成のボリュームを有しているかなどを把握している必要がある。副ボリュームの作成が終了すると、この報告を受けたSAN管理装置は、ストレージに対してボリュームペア分離を指示し、正ボリュームに関してはサーバからの通常の処理を続けさせたまま、副ボリュームからバックアップ装置へのデータのバックアップをバックアップサーバへ指示する。バックアップサーバは、副ボリュームのデータをSAN経由で読み出し、読み出したデータをバックアップ装置へ転送する。バックアップが終了すると、これがSAN管理装置へバックアップサーバから報告され、更に、SAN管理装置は、バックアップを依頼したアプリケーションへバックアップ終了を報告する。なお、ボリュームペアを分離する時間は、先のバックアップ時刻となる。また、バックアップデータのSAN上での転送先は、先のバックアップ装置のSAN上でのアドレスとなる。ここで、SAN管理装置とストレージとの間の制御情報の通信は、図1に示すSAN管理装置から、LAN、サーバ、SANを経由してストレージに通信することも可能であるが、図示しないSAN管理装置とストレージをLAN経由で直接接続して、この接続により先の制御情報を通信するようにしても良い。

上記説明では、SAN管理装置が中心となりバックアップの受付け、レプリカ作成・分離、バックアップ処理、バックアップの終了報告を制御したが、アプリケーションサーバ内のソフトとバックアップサーバのソフトが直接LANを経由して制御情報を交換することにより、SAN管理装置を利用せずに実現することも可能である(図6)。この場合、SAN管理装置を使用する場合と比較して、2つのサーバ内のソフトが相手のソフトとの連携が必要となるが、ここで説明したようなSAN管理装置を必要としなくて良く、比較的小規模なシステム向けと言える。

前述のバックアップでは、バックアップサーバを介してバックアップ装置へデータをバックアップしたが、バックアップサーバを経由せずに、ストレージ内の副ボリュームからSAN経由で直接バックアップ装置へデータを転送することでバックアップするように制御することも可能である（ダイレクトバックアップ）。これは、SAN管理装置を用いる場合は、SAN管理装置がレプリカ作成・分離を認識した後、ストレージに対して、副ボリュームのデータをバックアップ装置へ転送することを指示することで実現できる。この指示にはバックアップ装置のSAN上でのアドレス等が含まれることになる。

なお、上記バックアップでは、アプリケーションが主体となってバックアップファイル、ボリュームを指定する形にしているが、更新が頻繁に行われ、毎日や数時間毎のバックアップが必要なファイルやボリュームに対しては、管理装置やバックアップソフトウェアに定期的にバックアップすることを予め指定しておくことによって、アプリケーションへの負荷を減らすことが可能である。

次に、テープ共用バックアップ（図8）におけるSAN管理装置の働きの例を示す。LANフリーバックアップの場合、個々のサーバに関するデータのバックアップに関しては、上記のバックアップ動作とほぼ同じである。違いは、複数のサーバに関するデータのバックアップを行う必要が有るので、これら複数のサーバからのバックアップ処理による競合を制御する必要が有り、この競合を制御する機能がSAN管理装置に求められる。例えば、SAN管理装置は、予め定められたスケジュール等に従って複数のサーバにバックアップ指示を指示することによって、テープライブラリへのアクセスの集中を避ける機能等が必要となる。

SAN管理装置の動作例として図16に示すゾーニング（Zoning）機能を制御する例を示す。図16において、クラスタサーバは、ファブリック・スイッチ（Fabric Switch）を経由してストレージに接続される。ここで、ファブリックスイッチは、論理的に分割されており、即ち、複数のスイッチとしてみなされる。従って、ゾーニング1のスイッチのストレージ側出力先と、ゾーニング2またはゾー

ニング3のスイッチのストレージ側出力先を分離しておけば、ゾーニング1のスイッチに属しているクラスタサーバからは、ゾーニング2またはゾーニング3のスイッチをアクセスすることはできず、ゾーニング1に属するクラスタサーバからゾーニング2またはゾーニング3のスイッチのストレージ側出力先への不正アクセスを防ぐことができる。

この様な、スイッチへのゾーニングの設定は、ファブリック・スイッチと図示しないSAN管理装置を図示しないLAN等で接続し、SAN管理装置等からの指示でファブリックスイッチへ前述のようなゾーニングを設定することで行うことができる。SAN管理装置を用いない場合、ファブリックスイッチへ専用コンソール等を用いて設定することもできるが、クラスタサーバやストレージを追加・変更・削除する毎に前記専用コンソール等のところへ行きゾーニング用制御情報を設定する必要があり、非効率となる。SAN管理装置を用い、これから通信を用いて設定することで、その使い勝手が改善される。

以上、SAN管理装置の2、3の動作例を示したが、基本的には、色々なデータ処理機能を提供する際にSAN管理装置は、サーバやストレージから操作対象となるファイルやボリューム情報、操作タイミング、データ移動先、等を取得し、その情報に基づいて必要な装置にファイルやボリュームの処理（レプリカの作成、データコピー、レプリカの分離、バックアップコピー、リモートコピー、等）を、操作タイミング等の情報に基づいて指示する。各装置はその指示に基づき、処理を行い、処理結果をSAN管理装置へ返答する。それをサーバやストレージに必要な応じて報告する。必要に応じて、これら処理を依頼したクライアントへSAN管理装置から報告を返しても良い。

整理すると、SAN管理装置（図1の運用管理ソフトの端末）は、各アプリケーションサーバ上で動作しているアプリケーションから、統合ストレージシステム内のデータに関する処理の依頼を受け付ける第1ステップ（このステップは別途事前に定められたスケジュールに基づいてSAN管理装置が自発的にデータに関する要求を生成

するステップに置き換えることができる)、当該データに関する処理に必要な情報(処理対象データを特定する情報、操作時間、データ移動先、等)を得る第2ステップ、前記得た情報を元に、ストレージ、ネットワークスイッチ、サーバ等に存在する各種機能ソフト(レプリカの作成、データコピー、レプリカの分離、バックアップコピー、リモートコピー、等の実行ソフト)の起動順序や起動タイミング等の機能ソフト実行スケジュールを決定する第3ステップ(本ステップは各機能ソフトを連携させるステップとも云える)、決定した実行スケジュールに基づいて実際に機能ソフトを起動する第4ステップ、実行結果を各装置の機能ソフトから結果を得る第5ステップ(この第4ステップの結果は第3ステップの成果であるスケジュールに影響を与える場合もある)、第5ステップの結果を、データ処理を依頼したアプリケーションに報告する第6ステップを有するとも云える。なお、本ステップは便宜的に分けたステップであり、2つのステップを統合したり、一方、更に細かく分割して別ステップとしても問題ない。

以上説明したように、SAN管理装置は、複数の機能ソフトを連携させ運用する機能を持つので、個々の機能ソフトでは実現できなかった複雑な機能を容易に実現でき、統合ストレージシステムにおいてよりの確にデータ処理を行うことが可能になる。尚、複数の機能ソフトを連携させずに1つの大きなソフトを準備することで複雑な機能を実現することも可能であろうが、データ処理の種類毎に一々ソフトを開発する必要が発生し、融通の利かないシステムになる。

次に、大規模な計算機システムにおいて、ストレージシステム、ストレージエリアネットワーク技術がどのように使われるかを具体例を用いて以下説明する。図17は、最近拡大しつつあるインターネット・データセンタ(「iDC」とも略称する)の構成例を示している。インターネット・データセンタは、ISP(インターネット・サービス・プロバイダ)や各企業のWWWサーバを預かり(ハウジングという)、ネットワーク管理やサーバ運用管理を提供する。さらにウェブ・デザインやEC(エレクトリック・コマース)システムの構築、高度なセキュリティの追加など付加価値サービス

の提供も行う。インターネット・データセンタは、インターネットビジネスを行いたい企業が抱えるシステム要員やスキルの不足、サーバ設置場所やネットワーク網の整備などの問題を解決するソリューションを一括して提供している。

インターネット・データセンタの特徴として、高速ネットワーク回線などの高価な機器が共有されるため、データセンタの提供者からすれば低コストで多くの企業にサービスを提供できる。また、データセンタの利用者・利用企業からしても、バックアップ・保守などのわずらわしい作業から解放され、単独で運用するよりも低コストで済む。ただし、各企業が使うインターネット環境、アプリケーションソフトウェアを数多く稼動させるため、高速なインターネットバックボーン回線、多くの高性能サーバが必要となる。また、それらには当然のことながら、高信頼性、高セキュリティが必要とされる。これら環境では高速かつ高機能なストレージシステムの存在が必要不可欠である。

以下、インターネット・データセンタなどの大規模システムへのストレージエリアネットワーク技術の適用例を示していく。

図18は、インターネット・データセンタに大規模ストレージ・ストレージエリアネットワーク（SAN）を適用した場合の構成図の概略を示している。企業ごとに複数のサーバ計算機が存在するが、ディスク・テープなどのストレージは1台ないし23台のわずかな台数に集約され、サーバとディスク・テープ装置はファイバチャネルスイッチで相互結合される。SANが存在しない環境では、個々のサーバ計算機に個々のストレージ装置を接続する必要があるが、SANにより、ストレージ装置をすべての計算機から共有できるため、ストレージ装置を集約し、管理することが出来る。また、ストレージ装置の増設においても、ホスト計算機がオンラインの状態のまま増設が可能で、業務に影響を与えない。

また、バックアップの観点においても、SANによるストレージの集約化は有効な役割を果たす。ここで、図19はインターネット・データセンタ内SAN環境下での無停止バックアップの事例を示す構成図の概略である。ここでは複数企業のそれぞれ

のサーバ計算機、ストレージ、バックアップライブラリがストレージエリアネットワークにて相互結合される。そして、SAN上ストレージ・デバイスの管理、バックアップの運用を行うための管理ホストが存在する。各サーバ計算機のデータ、たとえばWWWサーバにおけるWebコンテンツ、アプリケーションサーバで使われるデータはSAN上のストレージに集約・格納されている。

バックアップは各ホスト計算機の都合によりさまざまな要求が考えられる。例えばホスト計算機へのアクセス負荷が低くなる時点で、すなわち深夜など、ディスクアクセス数が減る時間帯に毎日バックアップを取る、あるいは更新系トランザクション処理の多いホスト計算機の場合には、トランザクションの切れ目など、ホスト計算機がその時々に応じて任意のバックアップ開始時間を決定したいなどの場合が存在する。

管理ホストでは、ホスト計算機のそれら要求を受けて適切にバックアップ処理のマネジメントを行う。また、インターネット・データセンタは24時間連続運用が重要であるため、ホスト計算機の処理を中断することは避けなければならない、無停止バックアップが必須である。以下、バックアップ処理の一例を簡単に説明する。

例えば、各サーバ計算機が1日1回、あるタイミングでバックアップをとりたいという場合、管理ホストではそれぞれのサーバ計算機のバックアップ開始・終了のスケジューリングを行う。例えばA社のWWWサーバのバックアップを深夜0時から実行し、1時からB社のアプリケーションサーバのバックアップ、1時半からはA社のアプリケーションサーバのバックアップを行い、3時からB社のWWWサーバのバックアップを行うなどである。バックアップ処理に要する時間は各サーバの保有するデータ量などに依存するため、管理ホストは各サーバ計算機がどのくらいのデータをストレージに保有しているかなども管理し、その上でバックアップ所要時間を算出、スケジュールする。また、テープライブラリが複数のテープドライブを有していれば、複数のバックアップジョブを同時実行することも可能である。

A社のバックアップを深夜0時から実行する場合を例にとって処理の流れを説明する。深夜0時になったら、管理ホストはディスク装置内にあるA社WWWサーバのデ

ータのレプリカを作成する。そのために、ディスク装置内にある空きディスク（論理ボリューム）を探しだし、それをA社WWWサーバのレプリカボリュームとして割り当て、レプリカ作成をディスク装置に指示する。レプリカ作成の処理の流れについては、図5 a、図5 bの図示記載にて詳述されているとおりである。

引き続き、テープライブラリのテープドライブに、テープカートリッジを装填する。その後、レプリカボリュームからテープライブラリにデータをバックアップし始める。データのバックアップを実際に行うのは、A社サーバ計算機で行っても良いが、管理ホスト、あるいはディスク装置からテープライブラリにデータを直接転送するダイレクトバックアップ機能がサポートされている場合（ディスク装置、テープライブラリ、FCスイッチのどれかがサポートしていれば良い）、それを用いることも出来る。

そうした場合、サーバ計算機はバックアップが行われているかどうかについてまったく意識することなく、自動的にバックアップが取れていることになる。バックアップ処理が完了すれば、テープカートリッジをテープドライブから取り出し、ディスク装置のレプリカボリュームの使用を中止し、再び空きボリュームとし、次のバックアップ処理に移る。

この場合、SANによって、テープライブラリを共有、相互接続しているため、管理ホストなどの役割によってテープライブラリの利用スケジュールを適切に管理すれば、ホストが複数台あっても1台のテープライブラリですべてのバックアップをまかなうことが出来る。また、管理ホストのボリューム割り当てを適切に行うことにより、バックアップ処理の必要があるときだけレプリカボリュームを用意すればよくなり、常時各ボリュームにレプリカボリュームを用意しておく必要もなくなり、テープライブラリの台数、ボリュームの個数などを節約することが出来る。

次に、SANによるストレージ装置の共有はコスト削減のメリットが大きいですが、一方で複数企業のサーバが混在する環境下においては注意しなければならない点もある。その1つがセキュリティである。SANによりすべてのサーバ計算機がSAN上のすべてのストレージ装置にアクセスできてしまうため、A社のデータを、同じSAN上

にあるC社のサーバ計算機から覗けてしまう。次に、これらの問題を解決する手段の事例を説明していく。

図20はインターネット・データセンタにおいて複数企業のサーバ計算機・ストレージがSAN上に混在している環境を示している。図のように、A、B、C社がストレージを共有する環境下では、まずFCスイッチのゾーニング設定を行い、各社のサーバ計算機が、ストレージ装置の特定のパスのみしかアクセスできないように設定する。続いてディスク装置で、各パスに各社のサーバ計算機が使用するLUを割り当てる。例えば図のように、B社でLU1、LU2の2つの論理ユニットを使用する場合には、中央のパスにLU1、2をわりあて、C社でLU0を使う場合には、右側のパスにLU0を割り当てる。

さらに、同一パス上に複数のLUがあり、それを複数のサーバから共有する形になっているが、各サーバが共有したくない場合もあり得る。例えば、図20では、B社がLU1、2をアクセスできるパスを確保しているが、LU1はB社サーバの中でもある特定のサーバのみにアクセスを許可したい、という要求も有り得る。そういった場合には、LUNによるアクセス制限を実施する。B社の特定のサーバのWWNをディスク装置に登録し、LU1はこのWWNが登録されたサーバだけがアクセスできるように設定することが出来る。

これらゾーニング、パス割り当て、LUN単位のアクセス制限の設定は集中管理コンソールにて行う。管理コンソールにてFCスイッチのトポロジーを確認し、それをもとにゾーニングの設定をし、さらに各パスに必要なだけのLUをマップさせて、各社が使用できるLUを登録する。さらに同一パス内で相互アクセスを許可したくないLUについては、集中管理コンソールはアクセスを許可するホスト計算機のWWNを取得し、それをディスク装置に設定、LUN単位のアクセス制限を実施する。

次に、SANと各種ストレージからなる統合ストレージシステムを用いた計算機システムについての応用例を説明する。近年企業の合併・統合が増えている。それにより、企業間の計算機システムも統合する必要性が出てくる。

図21は複数の企業の計算機システムが相互結合された大規模計算機システムの事例を示す。企業間のホスト計算機はインターネットで接続され、データの相互利用などが図られる。また、ストレージエリアネットワークの導入により、ストレージについても各企業内のストレージ同士が公衆回線網、あるいは専用線で相互接続される構成がとれる。

計算機システムの運用の観点において重要な点は、データの統合にある。通常、各企業で用いられるアプリケーション・データベースは異なっており、装置が単純に相互接続されただけではデータを直接相互利用することはできない。そこで、一般には複数のデータベースからのデータを集約・統合した、新規のデータベースを構築していく作業が必要となる。

図21において、企業A、Bとでそれぞれ勘定処理などのトランザクション処理を行う基幹系データベース、基幹系データベースのデータを利用しオフラインでの解析処理を行う情報系データベースとを持つ。この事例では、企業A、企業Bの基幹系データベースのデータを統合し、各種業務用のデータマートを作る。場合によっては、一旦単一の大規模なデータウェアハウスを作成し、そこから各種業務アプリケーション向けの小規模なデータマートを作る流れもある。ストレージエリアネットワークによるストレージ同士が接続された環境が存在しない場合には、データベースの統合の際、ホストコンピュータ・ネットワーク経由でデータを移動する必要がある。通常、企業間で共有したいデータベースは大容量のものが多いため、データ転送に大変時間がかかる。

図21の事例では、ストレージのリモートコピー機能を用い、企業Bのレプリカを作成する。1日1回、あるいは1週間に1回などの周期で、一旦レプリカボリュームをスプリットし、スプリットされたレプリカボリュームの内容をレプリケーションサーバが読み出し、各種データマートを作成していく。レプリケーションサーバはデータマートを利用する各種情報系DBMSとは別に存在する。ストレージエリアネットワークでストレージが相互結合され、ストレージのリモートコピー機能を用いること

により、ホストに一切負荷をかけることなくデータベースのレプリカが作成できる。

またデータマートを作成するレプリケーションサーバと、情報系DBMSとがそれぞれ別のホスト計算機で実現できることから、データマート作成の処理が、基幹系DB、情報系DBの業務に影響を与えずにすむ。

本発明によれば、SANを用いたストレージシステムにおける構成要素又は機能の連携を強化することで統合ストレージシステムを構築でき、図3に示す種々の機能を全て奏させることができる。

さらに、統合ストレージシステムを、インターネットに接続して大容量のデータ保存とその利用を図るインターネットデータセンタに適用することによって、インターネット情報サービスをコスト並びに量と質の両面で効率良く且つタイムリーに提供することができる。

006290 05090960

クレーム

1. 複数のクライアントコンピュータ、複数の各種サーバ、データを保存する複数の各種ストレージ、前記コンピュータと前記サーバを接続するローカルエリアネットワーク（LAN）、前記サーバと前記ストレージの間に介在するストレージエリアネットワーク（SAN）、を備えた計算機システムであって、

前記SANは、ファイバチャネルスイッチ（FCスイッチ）によって前記サーバと前記ストレージを任意に相互接続し得る回線網を形成し、

前記各種ストレージの論理ボリューム、データ配置、障害監視を含むストレージ管理と、前記FCスイッチの設定管理と、前記ストレージのデータのバックアップ運用と、を行う管理運用ソフトウェアを備えた端末を有することを特徴とする計算機システム。

2. 請求項1に記載の計算機システムにおいて、前記SANは他の計算機システム内のSANとワイドエリアネットワーク（WAN）を通して接続されることを特徴とする計算機システム。

3. 請求項1に記載の計算機システムにおいて、ストレージの正ボリューム内のデータをバックアップ装置に無停止バックアップする場合、当該ストレージの内部機能によって正ボリュームに対応する副ボリュームを作成して前記正ボリュームから副ボリュームにコピーを作成し、作成されたコピーを前記LANを経由することなく前記SAN経由で前記バックアップ装置に転送してバックアップすることを特徴とする計算機システム。

4. 複数のクライアントコンピュータ、複数の各種サーバ、データを保存する複数の各種ストレージ、前記コンピュータと前記サーバを接続するローカルエリアネットワーク（LAN）、前記サーバと前記ストレージの間に介在するストレージエリアネット

ワーク（SAN）、を備えた計算機システムであって、

前記SANは、ファイバチャネルスイッチ（FCスイッチ）によって前記サーバと前記ストレージを任意に相互接続し得る回線網を形成し、

ストレージ内のデータをバックアップ装置に無停止バックアップする場合、当該ストレージは、サーバからのボリューム分割の指示を受ける機能を有し、当該指示時点で正ボリュームのデータがそのまま副ボリュームに保存されているように見せかける機能を有すると共に、前記副ボリュームからバックアップ装置にバックアップする機能を有することを特徴とする計算機システム。

5. サーバ、前記サーバのデータを記録するストレージ、前記サーバと前記ストレージを接続するネットワーク、前記ネットワークに接続されて前記データのバックアップを取るバックアップ装置、から構成されるシステムの管理方法であって、

処理対象であるデータを識別する情報を入手する第1ステップ、

前記情報が示すデータの処理内容を入手する第2ステップ、

前記情報が示すデータを保持する前記ストレージに対して、前記処理内容を指示する第3ステップ、および

前記情報が示すデータが処理された結果を前記ストレージから受ける第4ステップを有することを特徴とするシステムの管理方法。

6. 請求項5に記載のシステムの管理方法において、前記処理内容は、前記ストレージから前記バックアップ装置へ前記データを転送することであることを特徴とするシステムの管理方法。

7. 請求項5に記載のシステムの管理方法において、前記処理内容は、前記情報で指示されるデータの複製を作成することと、前記作成した複製データを前記バックアップ装置へ転送することであることを特徴とするシステムの管理方法。

8. 請求項5に記載のシステムの管理方法において、更に、当該処理内容を実行するタイミングを入手する第5ステップと、前記タイミングに従い前記第3ステップの実行タイミングを制御する第6ステップとを有することを特徴とするシステムの管理方法。

9. 請求項5に記載のシステムの管理方法において、前記システムの前記サーバはインターネットに接続され、前記データを前記インターネットへ送出するものであることを特徴とするシステムの管理方法。

10. 請求項6に記載のシステムの管理方法において、前記システムの前記サーバはインターネットに接続され、前記データを前記インターネットへ送出するものであることを特徴とするシステムの管理方法。

11. 請求項7に記載のシステムの管理方法において、前記システムの前記サーバはインターネットに接続され、前記データを前記インターネットへ送出するものであることを特徴とするシステムの管理方法。

12. 請求項8に記載のシステムの管理方法において、前記システムの前記サーバはインターネットに接続され、前記データを前記インターネットへ送出するものであることを特徴とするシステムの管理方法。

アブストラクト

ストレージエリアネットワーク (SAN) を用いたストレージシステムにおける構成要素又は機能の連携を強化することで統合ストレージシステムを構築するために、複数のクライアントコンピュータ、複数の各種サーバ、データを保存する複数の各種ストレージ、コンピュータとサーバを接続するローカルエリアネットワーク (LAN)、サーバと前記ストレージの間に介在する SAN、を備えた計算機システムにおいて、SANは、ファイバチャネルスイッチ (FCスイッチ) によってサーバとストレージを任意に相互接続し得る回線網を形成し、各種ストレージの論理ボリューム、データ配置、障害監視を含むストレージ管理と、前記 FCスイッチの設定管理と、前記ストレージのデータのバックアップ運用とを行う管理運用ソフトウェアを備えた端末を有する。

006290.05090560

Fig. 6

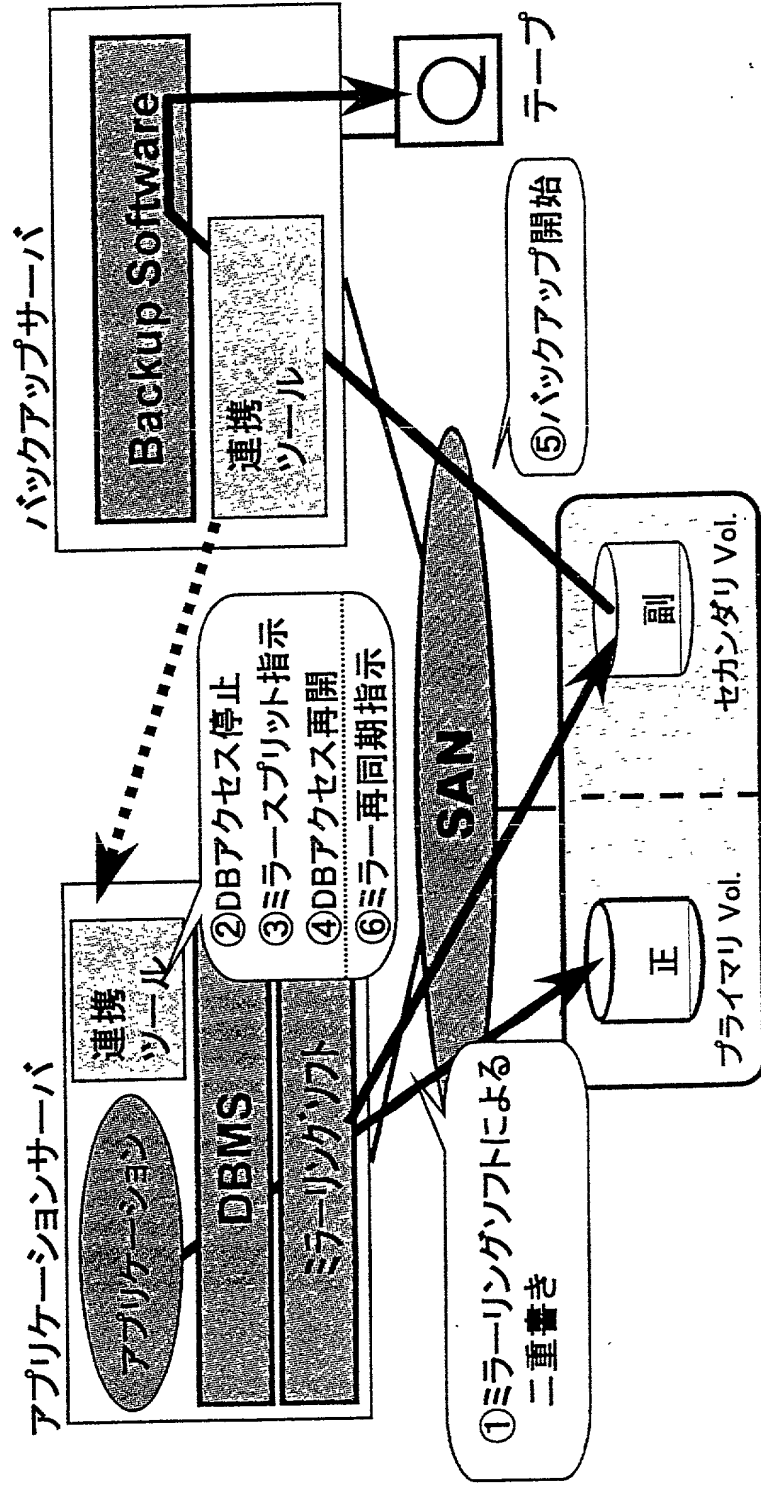


Fig. 7

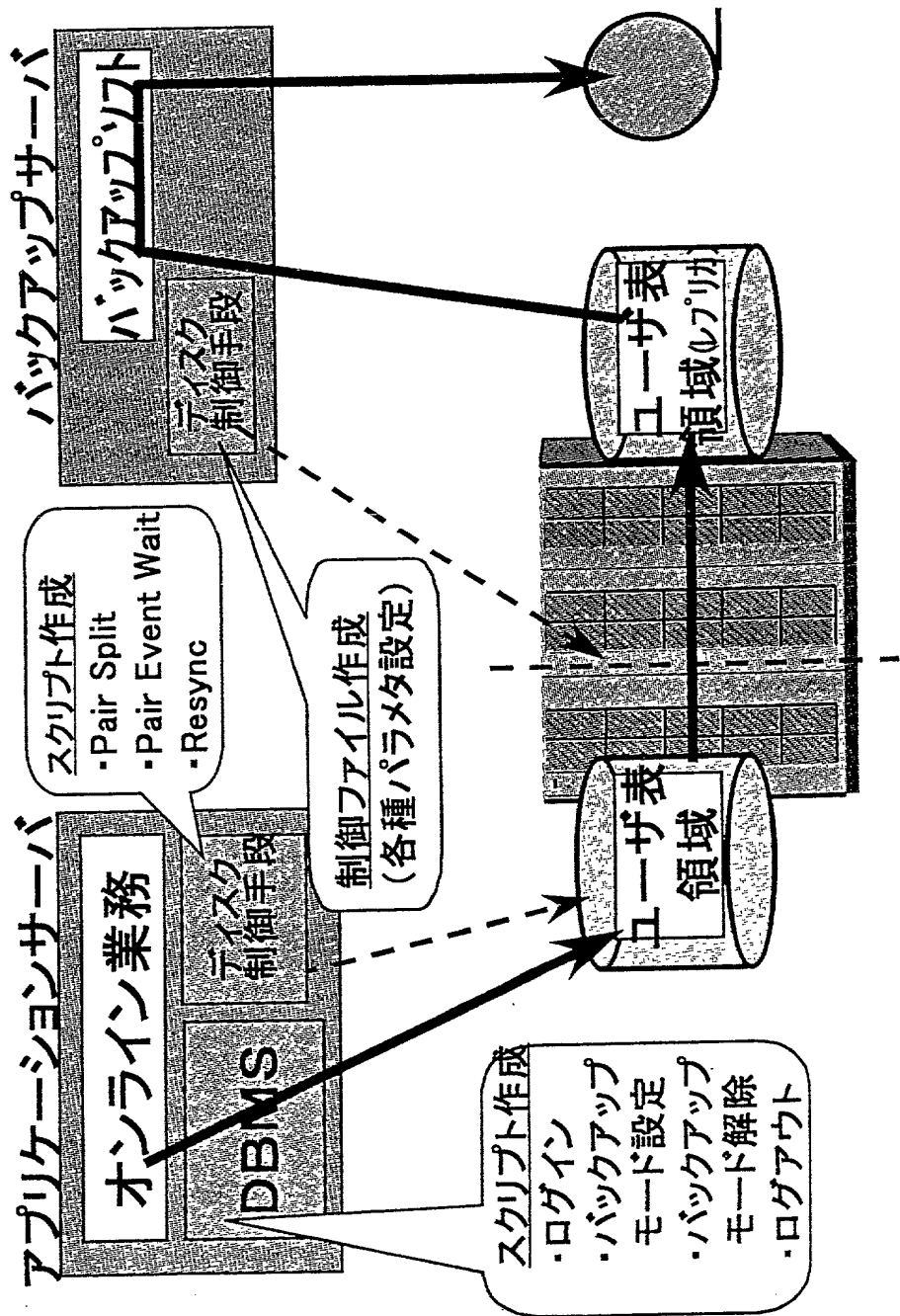


Fig. 8

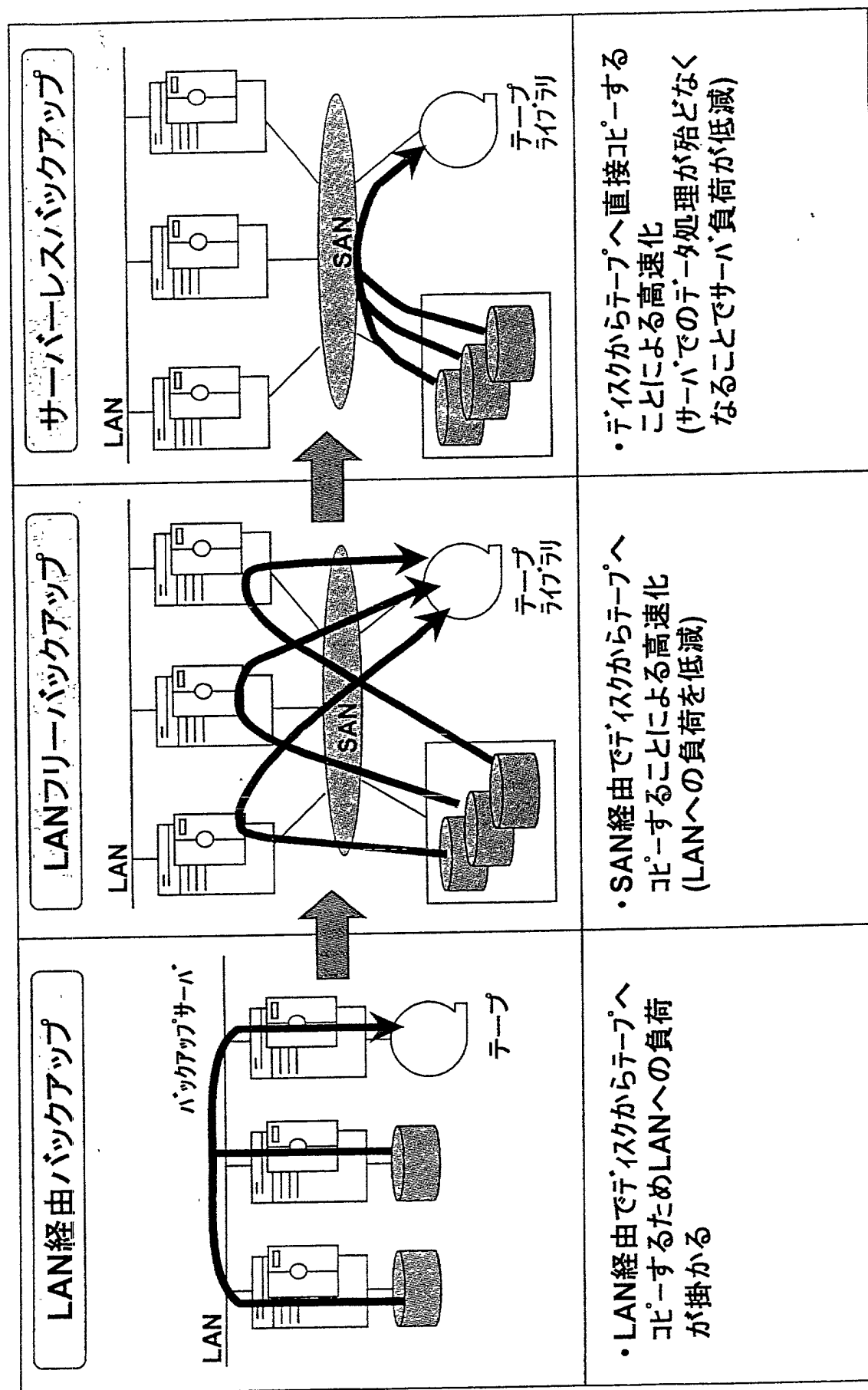


Fig. 9

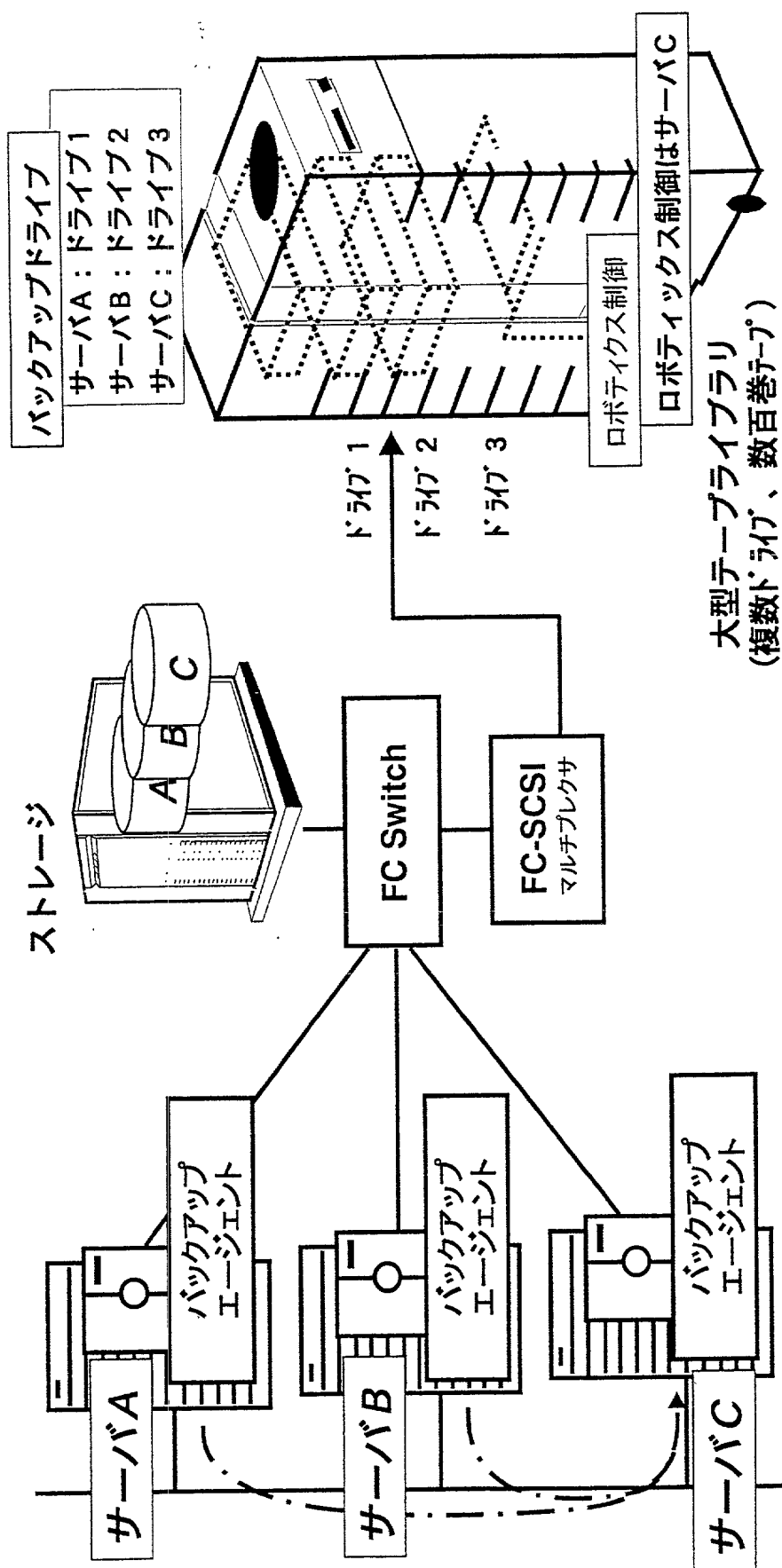


Fig. 10

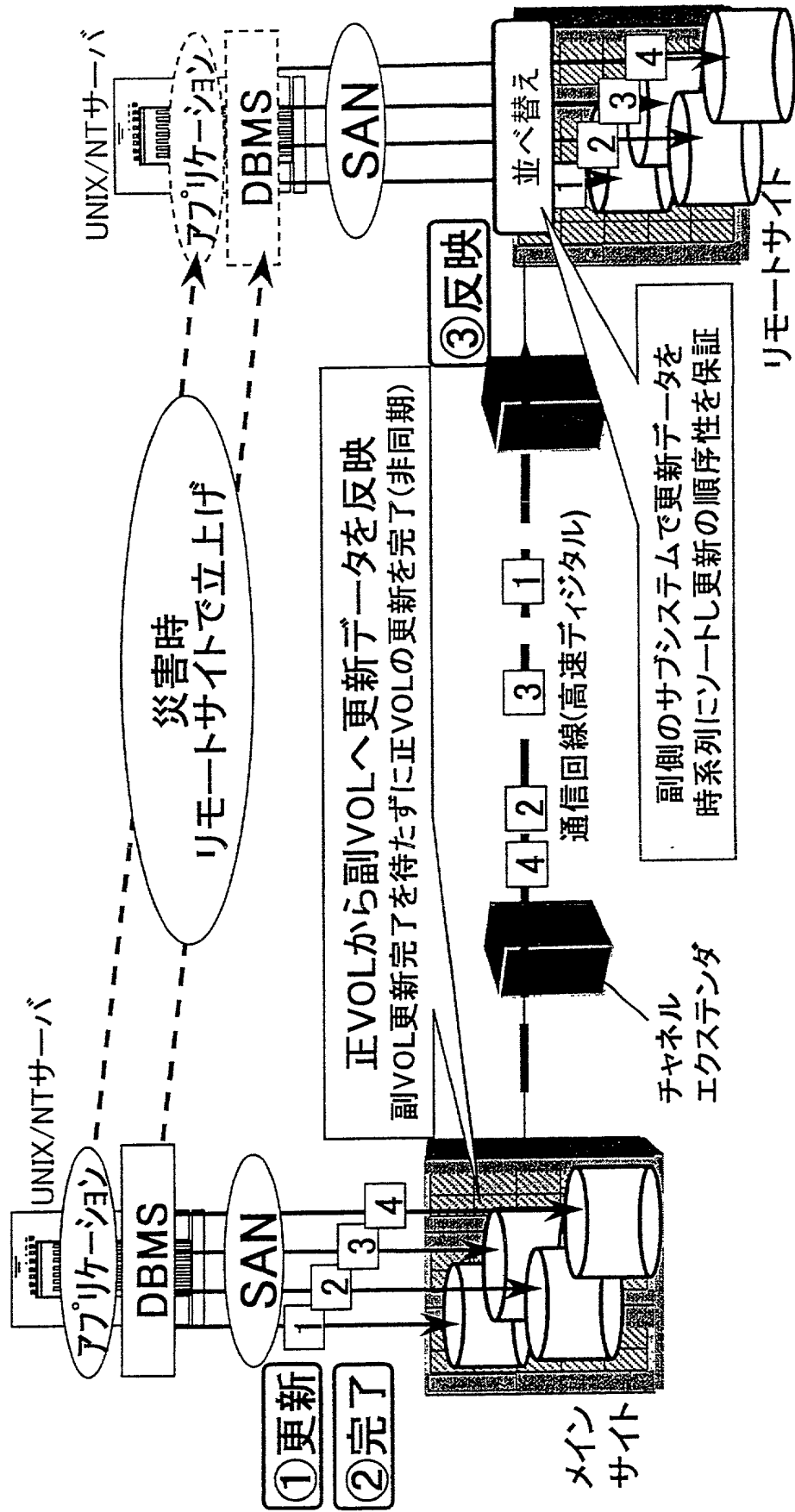


Fig. 11

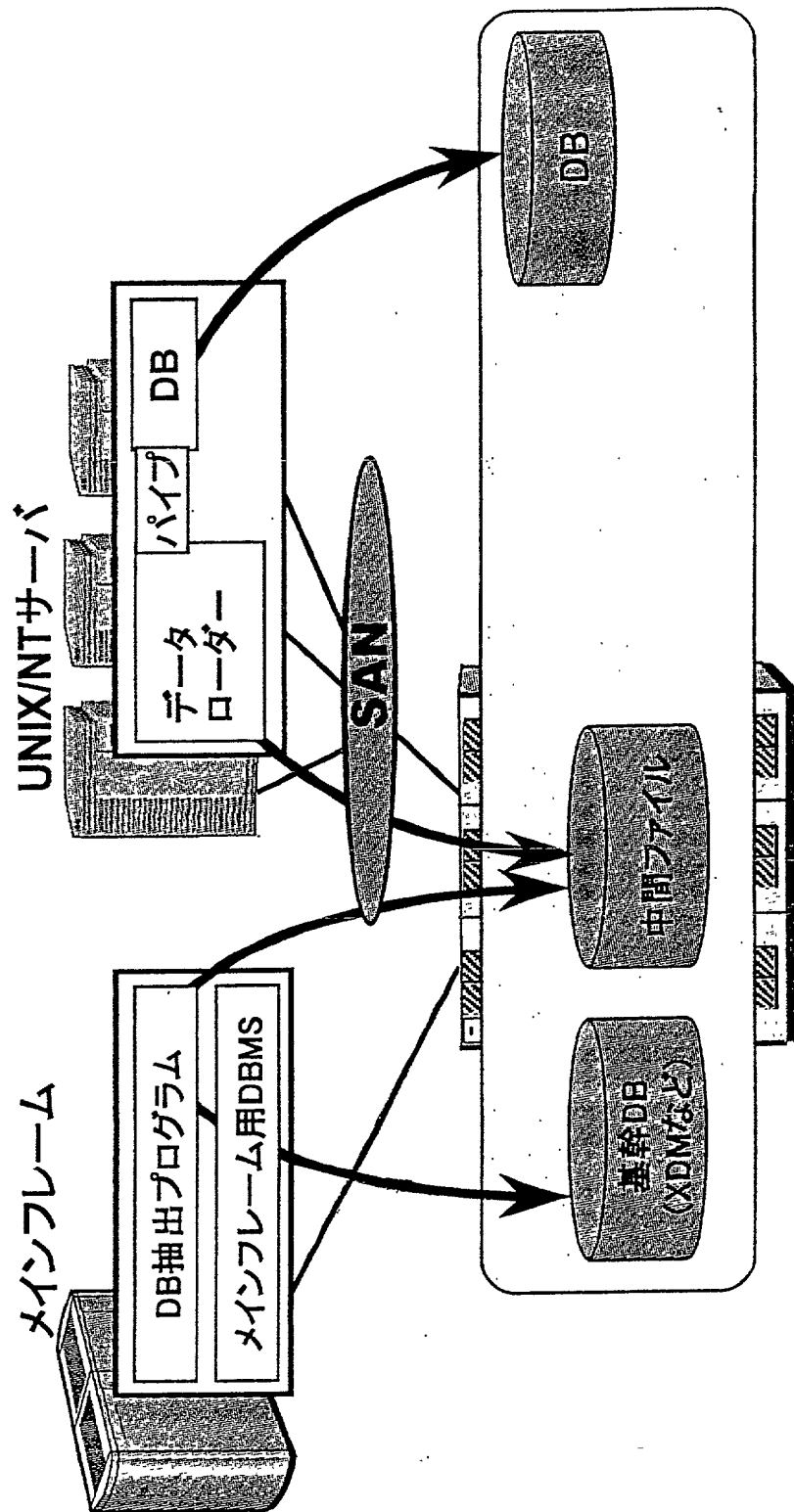


Fig. 12

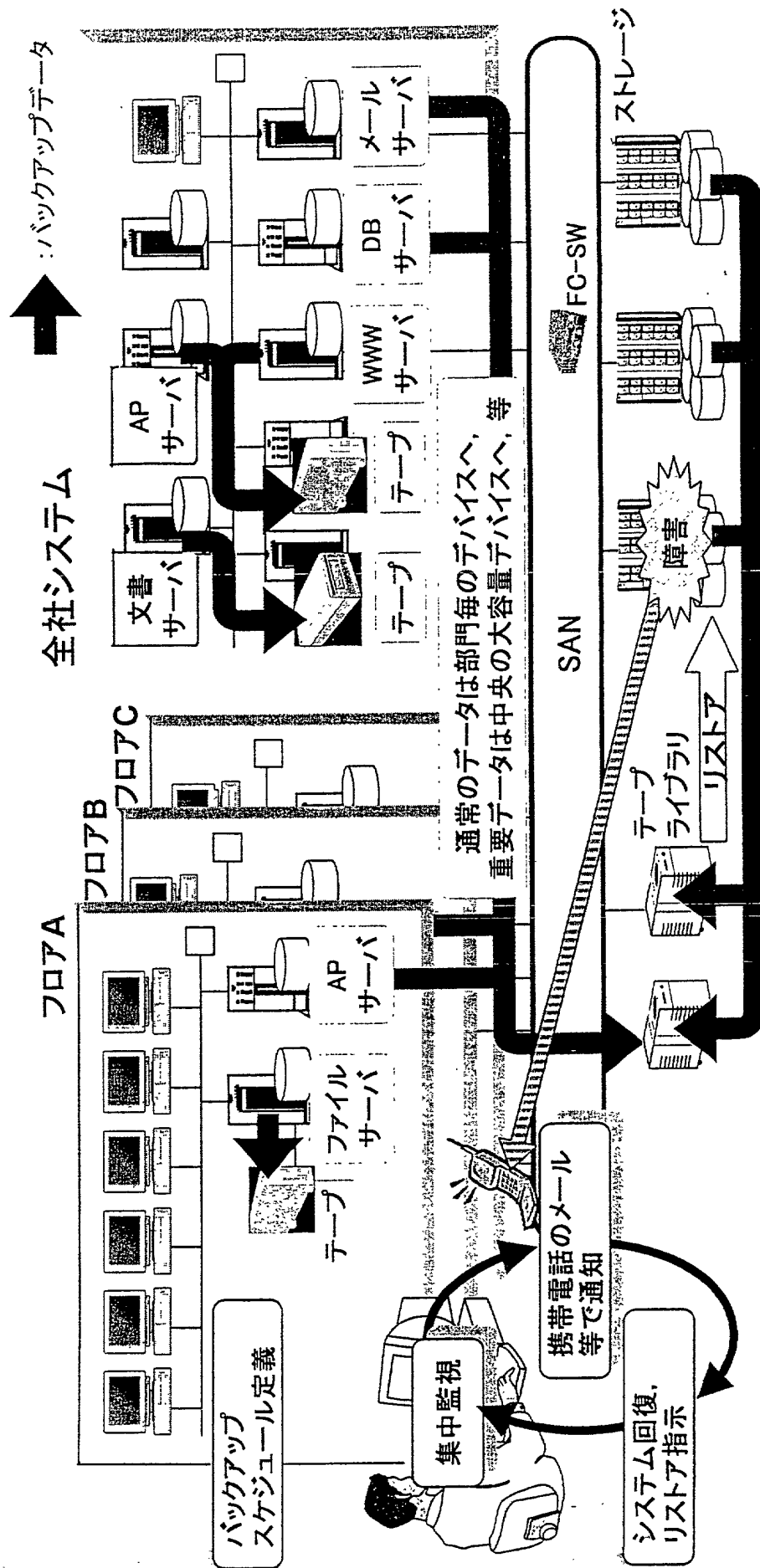


Fig. 13

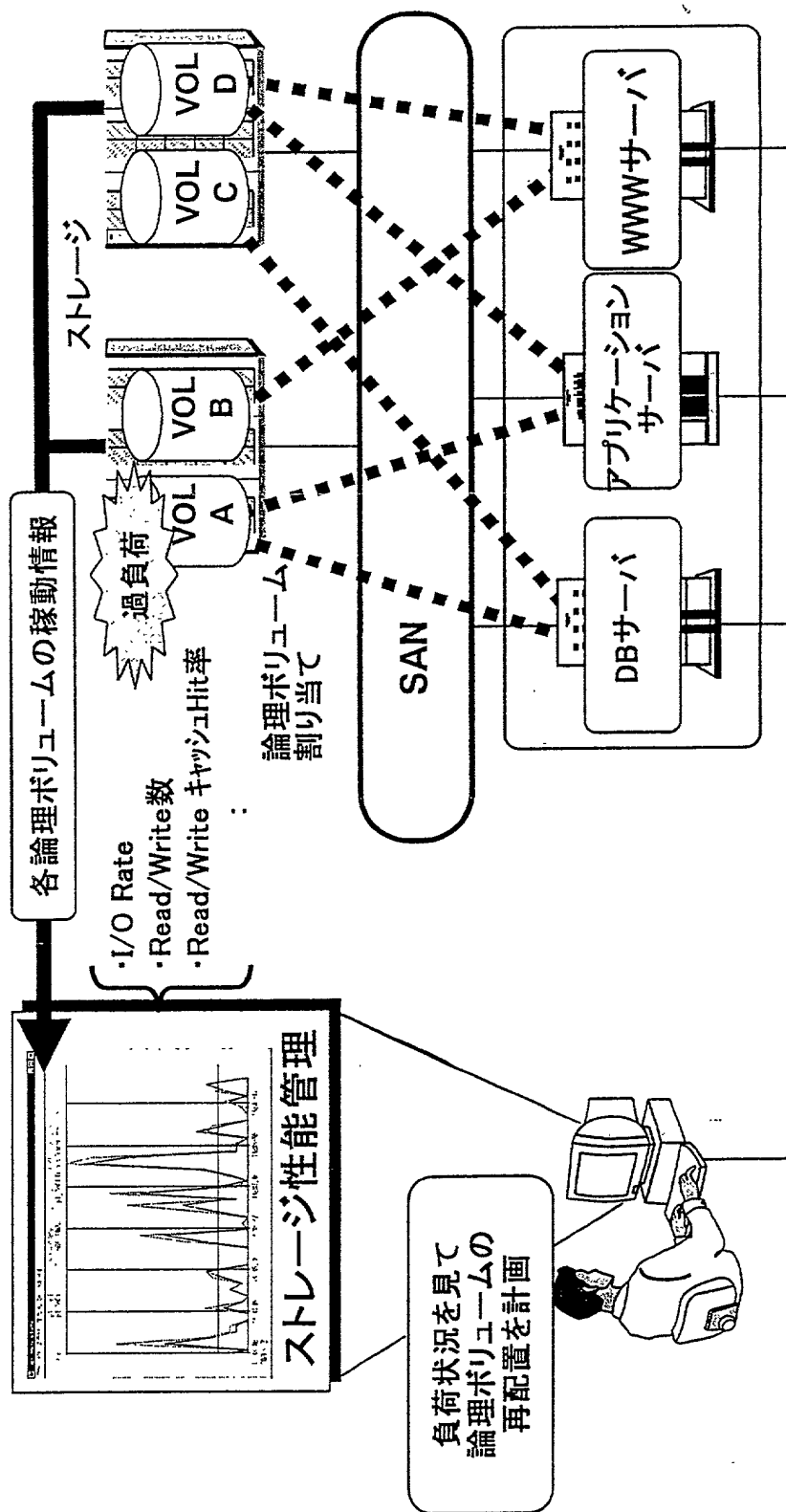
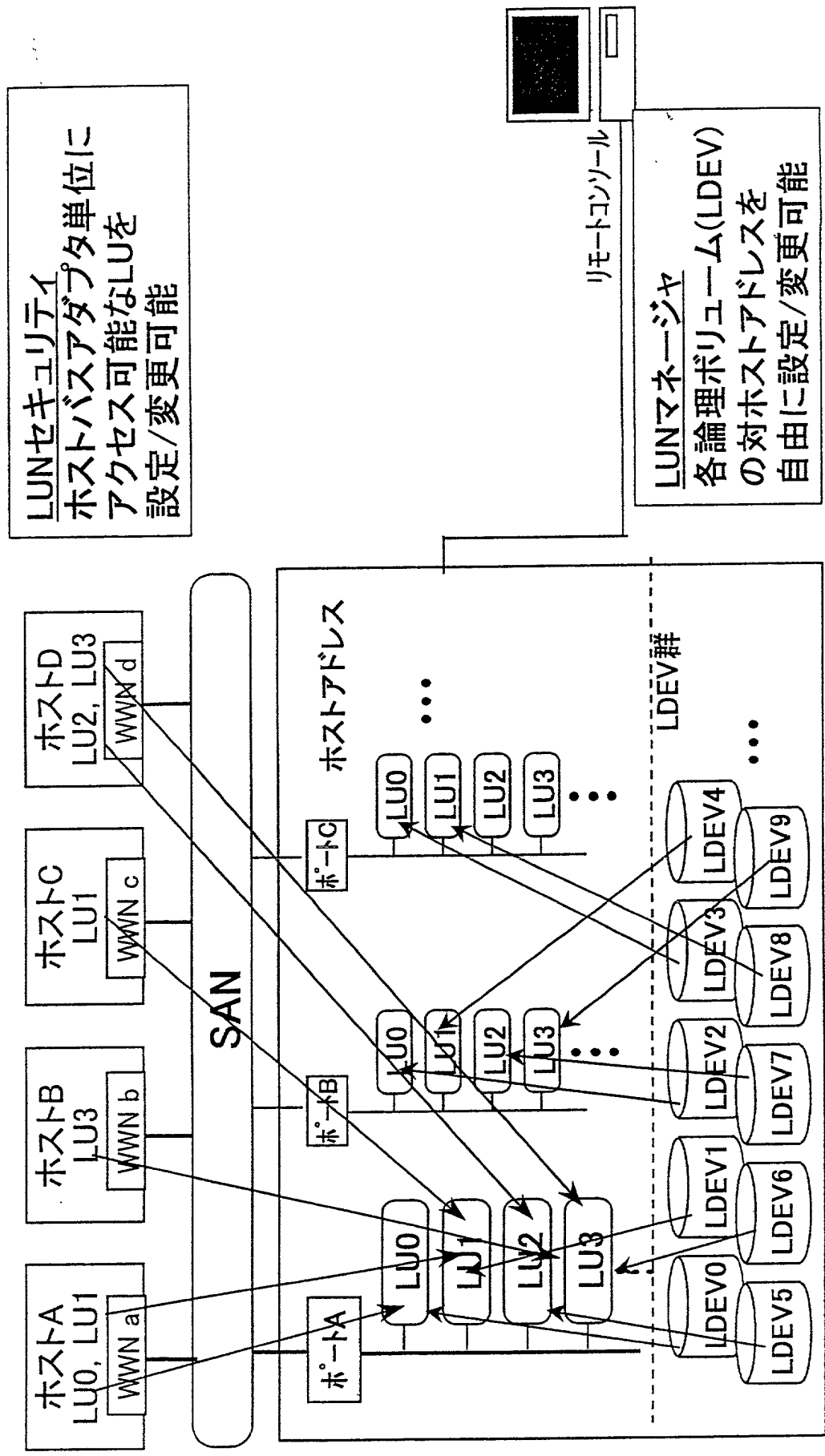


Fig. 14



LUN: Logical Unit Number、LDEV: Logical Device 管理論理ボリューム

Fig. 15

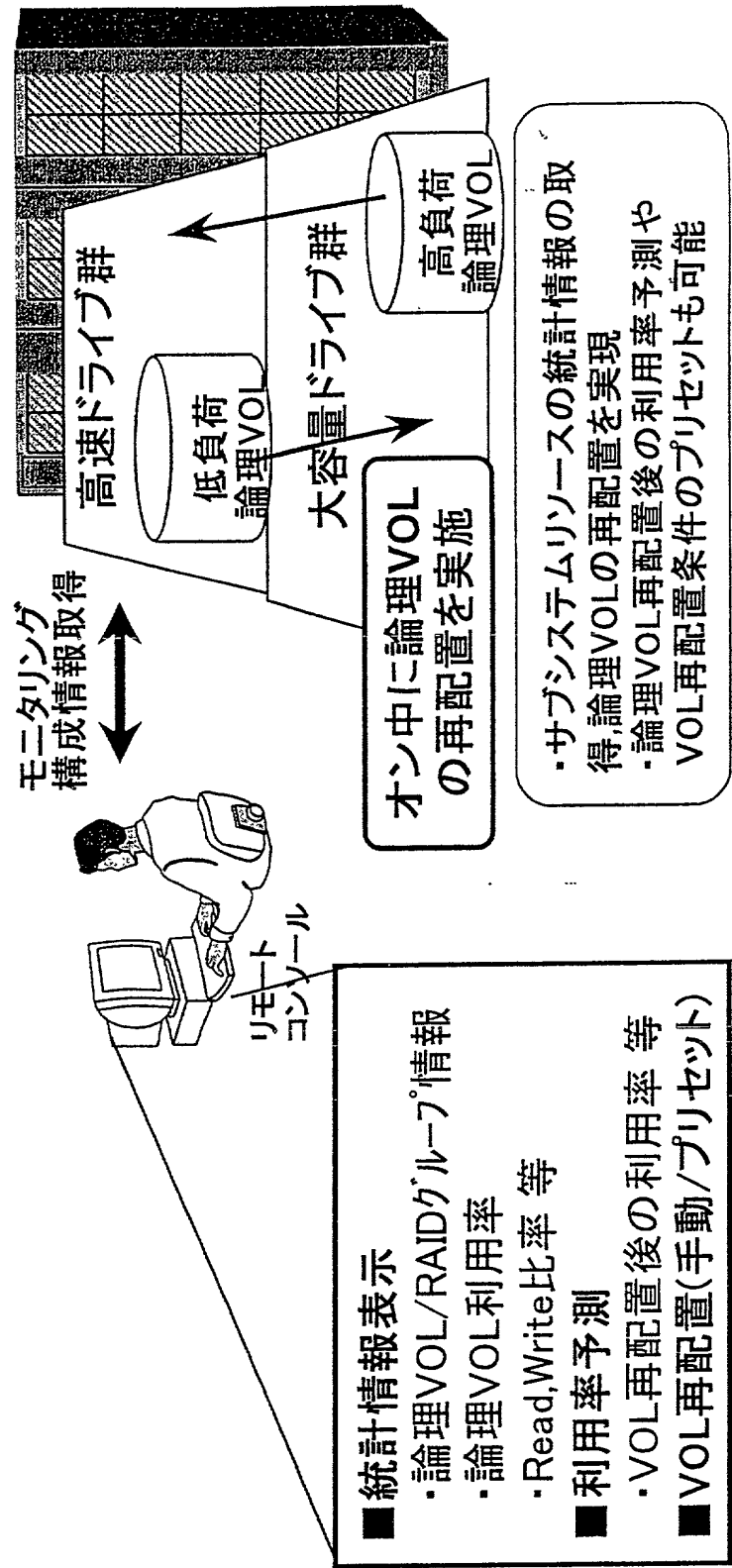


Fig. 16

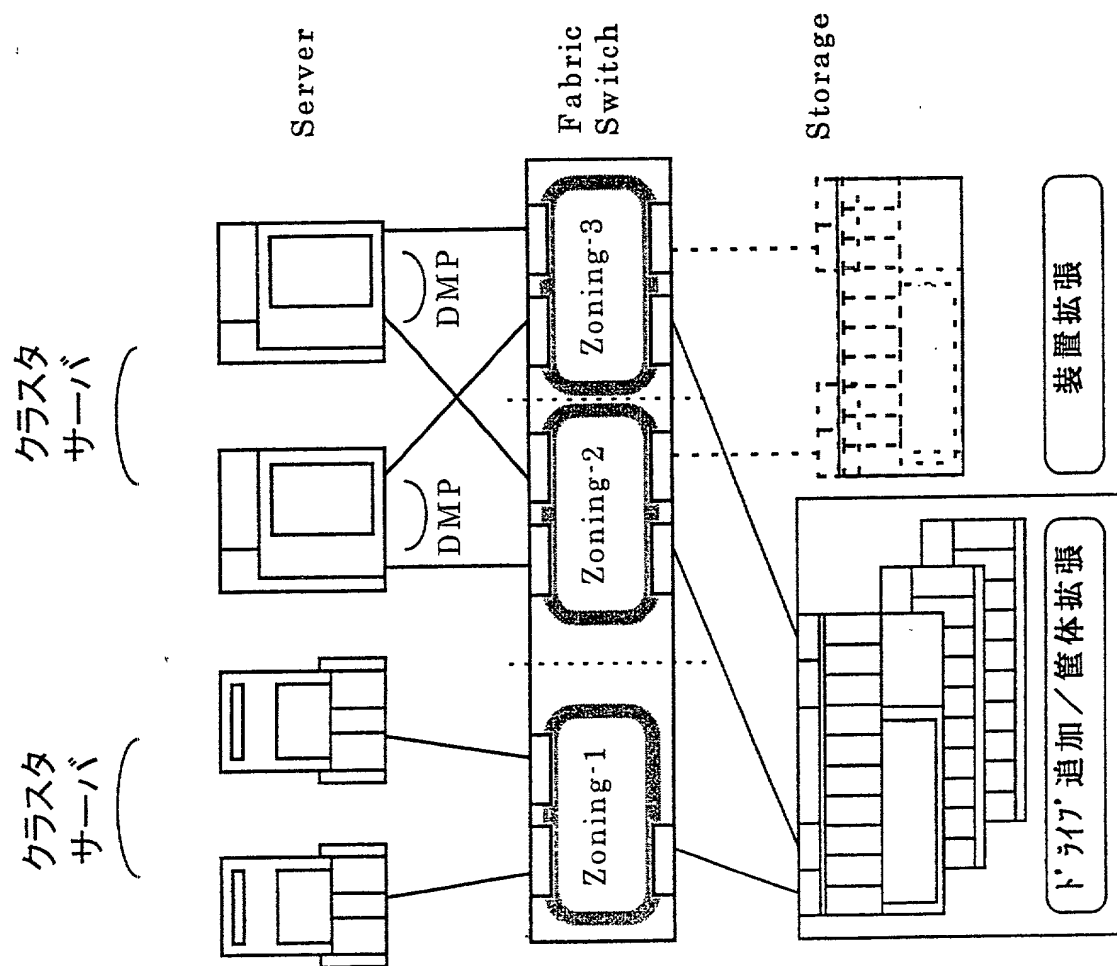


Fig. 17

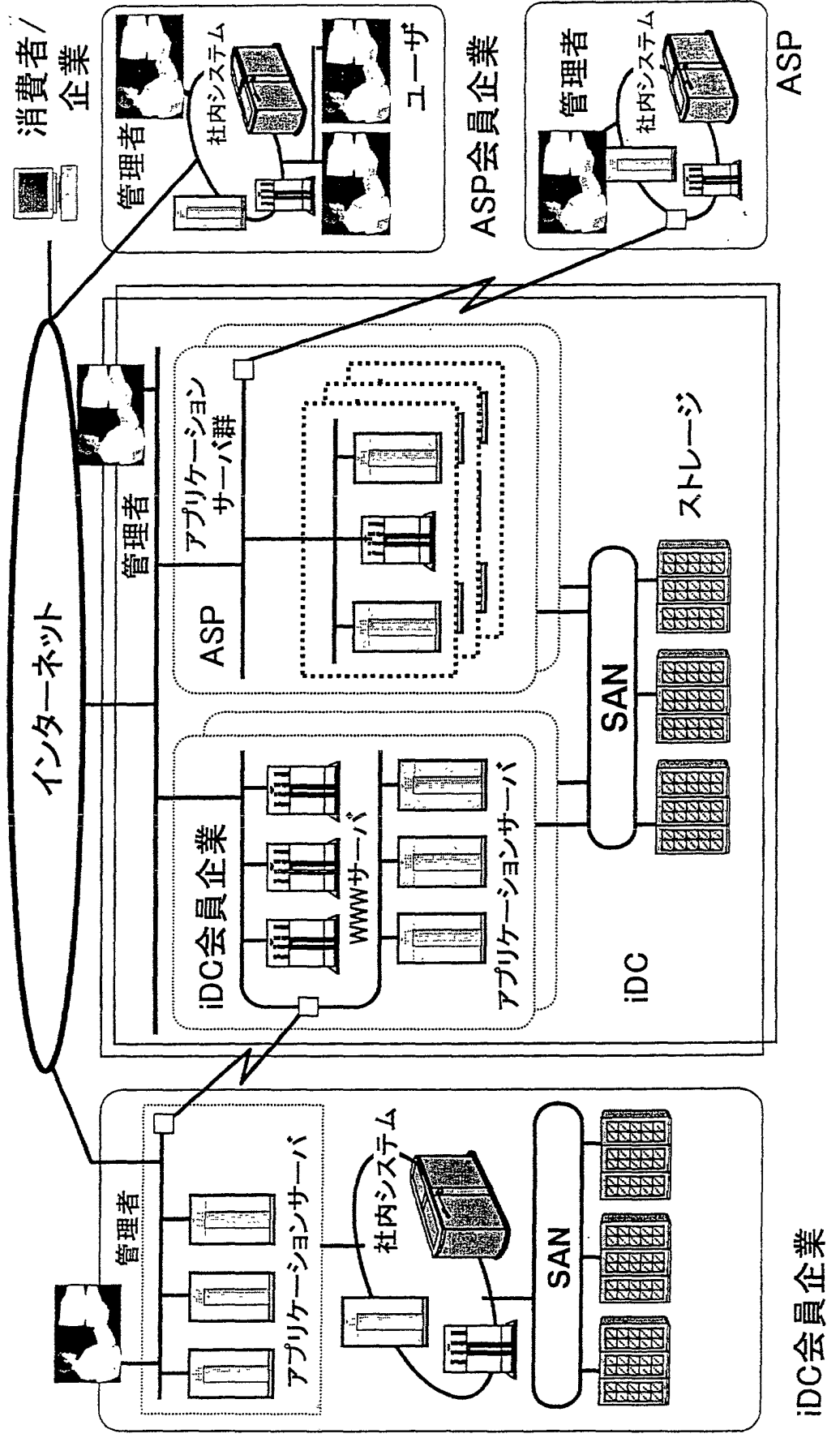


Fig. 18

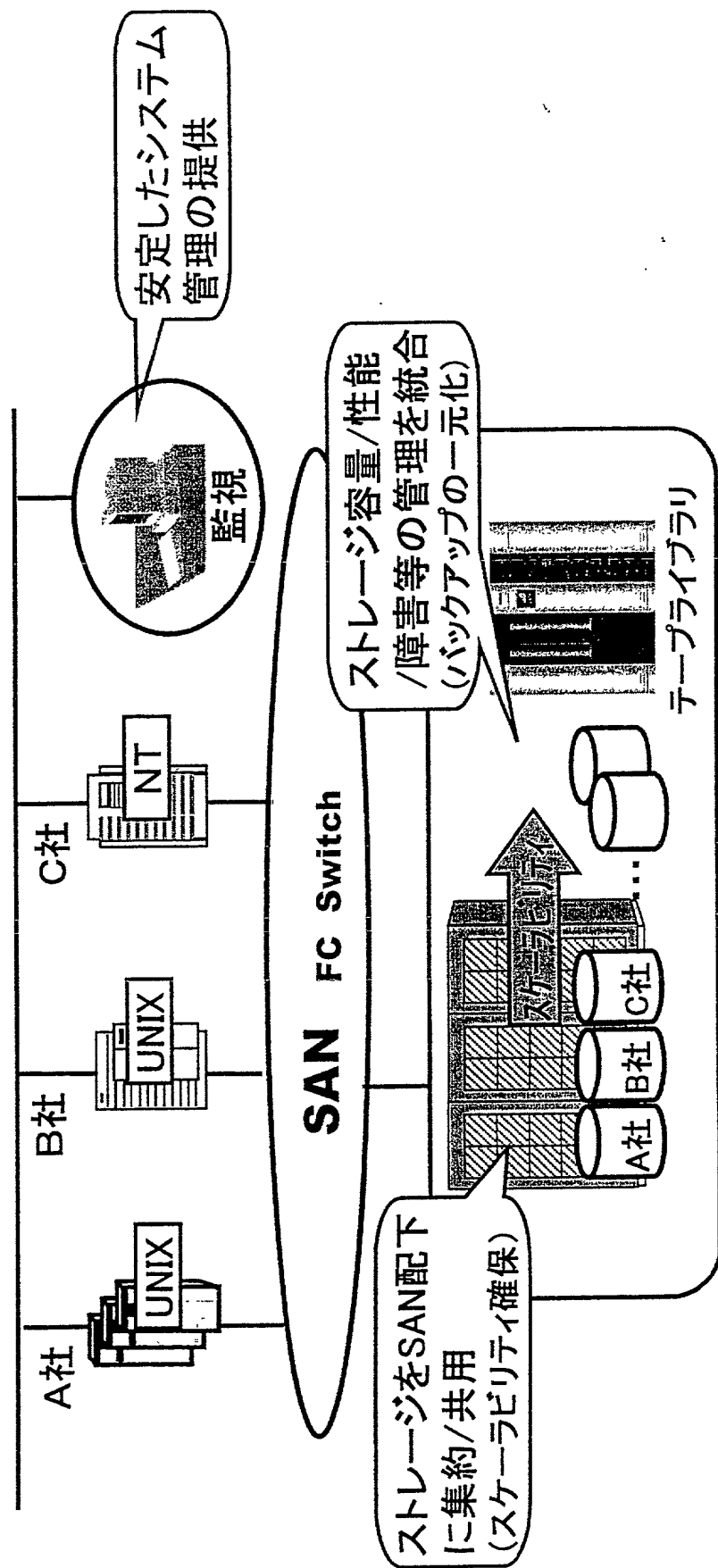


Fig. 19

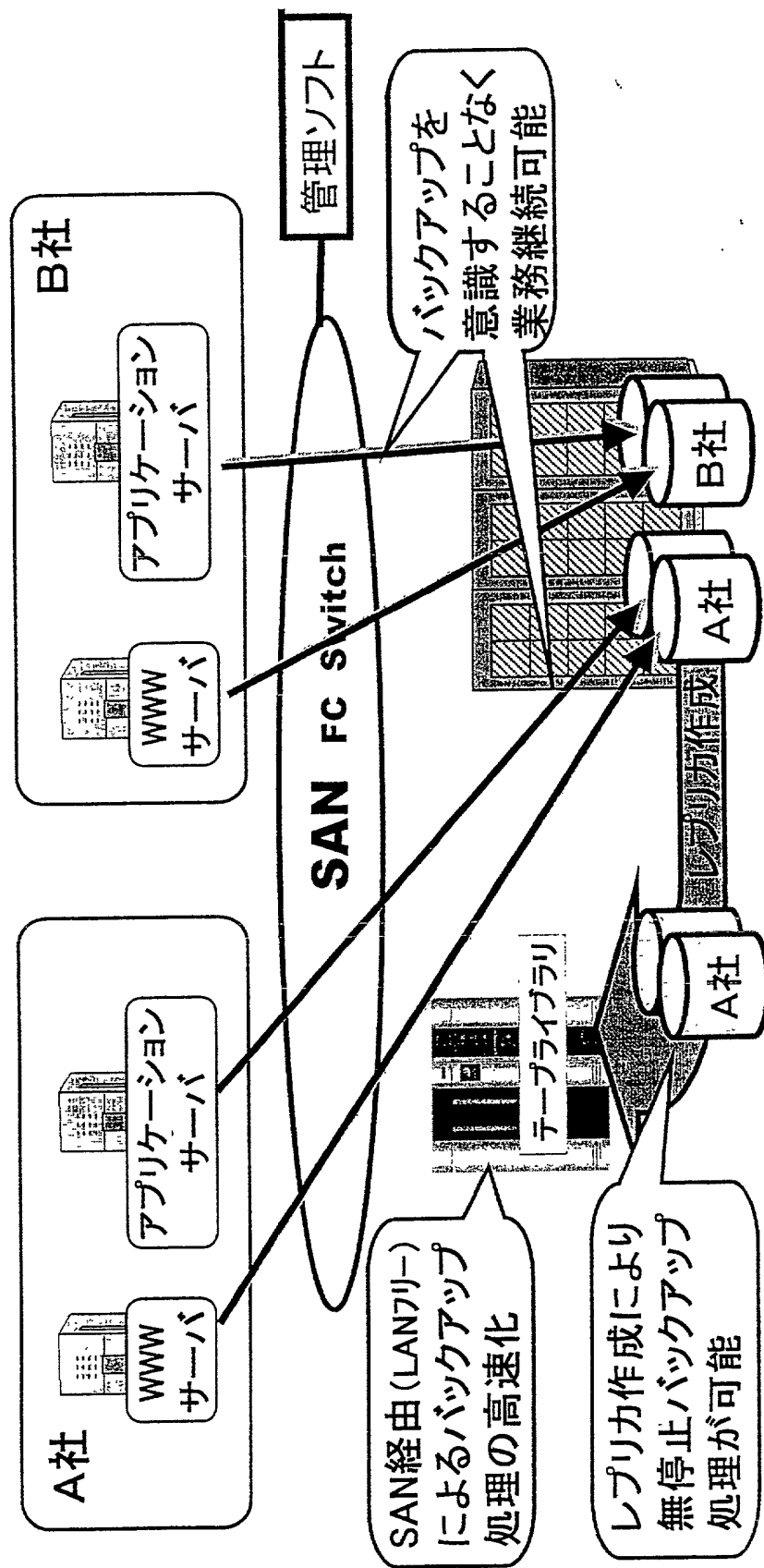


Fig. 20

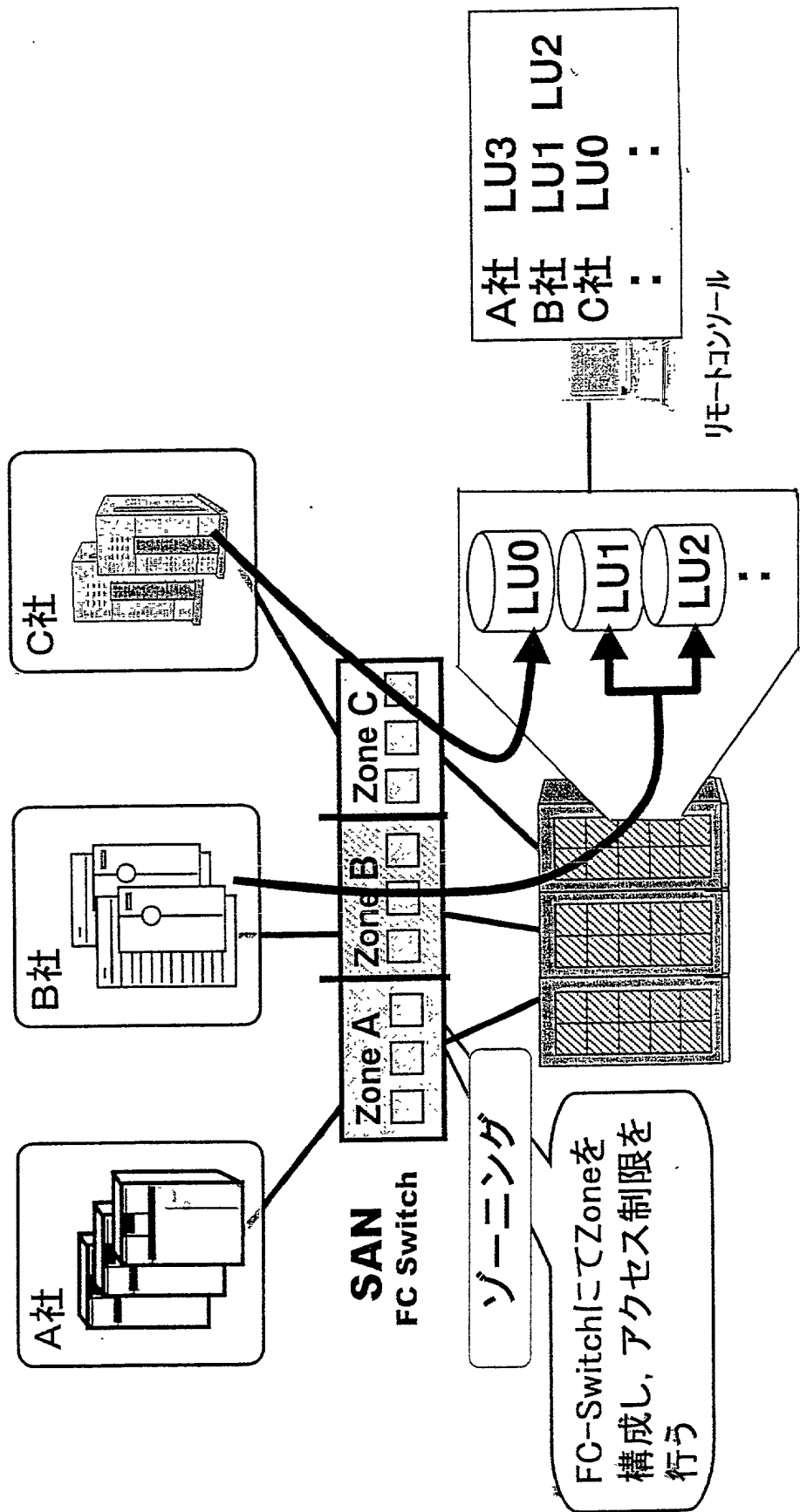


Fig. 21

